

# **Global Proteomic Approaches to Identify Potential Protein Biomarkers for COPD**

Prepared for  
**Philip Morris U.S.A.**  
August 27, 2003

Joel G. Pounds  
Joshua N. Adkins, K Monica Lee,  
Richard D. Smith, Kenneth Auberry, Nikola Tolic

Battelle Northwest / Pacific Northwest National Laboratory

**Battelle**

**Philip Morris / TNW / PNNL**

PM3001148253

# Outline

## Introduction

Biomarkers: Trends in risk assessment

BALF Proteomics - Current State-of-the-Art

Battelle / PNNL Biomarker Approach

Proteomic analysis of mouse BAL fluid

Objectives, design & approach

*Can MS be used to define proteins in BALF?*

*Can MS identify Strain- and exposure-dependent differences in protein composition?*

Future directions

Philip Morris/TNW/PNNL

Battelle

# Biomarkers: Trends in Risk Assessment

## Present

Major emphasis on hazard identification

Typical tests for detection of toxicity

Empirical cross species extrapolation

Dose metrics often unrelated to assessment of biological response

Methods for hazard identification in animals often different from man

Predictive value of animal data questioned

## Future

Increasing emphasis on hazard characterization

Exploration of mechanisms at cellular and molecular level

Mechanism-based risk assessment

Selection of dose metrics based on integration with biological response

Methods for safety assessment in human more in line with those in animals

Better inter-species correlation between animals and humans

# Why Bronchoalveolar Lavage Fluid (BALF)?

## Advantages:

- Reasonably assessable and closely related to target cells and tissues for lung diseases
- History of use for clinical and experimental studies
- Less temporally dynamic than cellular proteins
- Provide opportunity for direct cross-species comparisons

## Disadvantages:

- Invasive - not appropriate for routine measurement in humans
- Proteins may be diluted by lavage fluid

# Competing proteomic technologies MALDI, SELDI, 2-D gel, QTOF, LC-MS/MS, etc.

## Advantages

- Off the shelf hardware
- Relatively inexpensive
- Known technology
- Several 2-d PAGE databases are available



Battelle

## Limitations

- Sensitivity >100 fmole
- Relatively slow
- Reproducibility
- Requires spot mapping software
- Very basic polypeptides difficult to resolve
- Not quantitative

Philip Morris/TNW/PNNL

# Fourier Transform Ion Cyclotron Resonance Facility

## Sample Preparation

- Stable isotope labeling (ICAT, PhIAT), protein expression

## High-Efficiency Separations

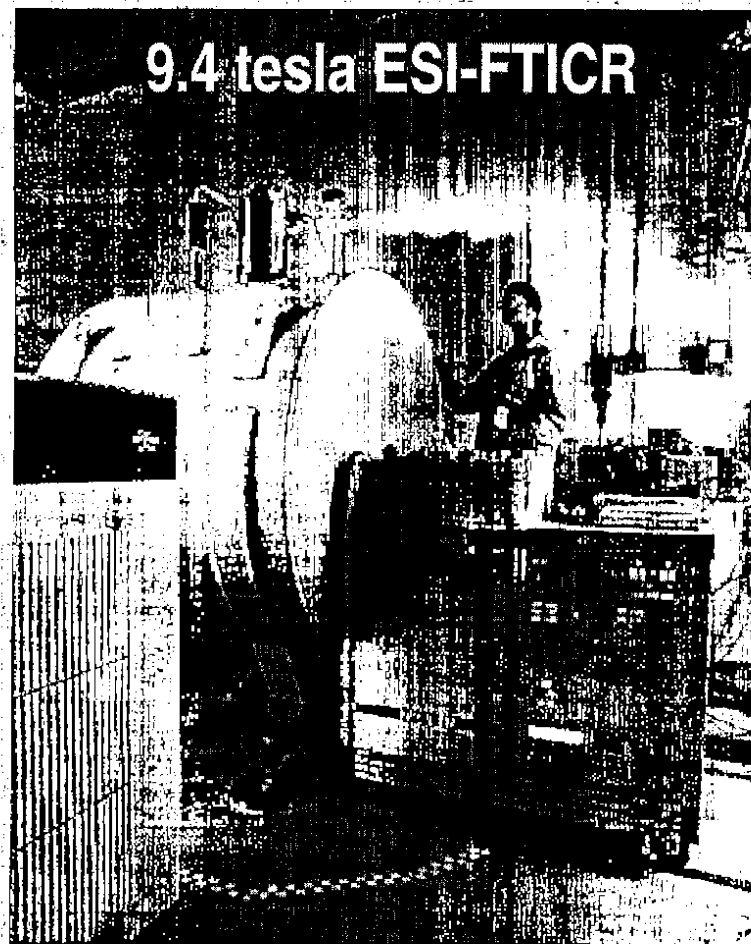
- LC separations at ~ 1000 psi; 10,000 peak capacity

## Mass spectrometry

- FTICR (3T, 7T, 9.4T, 11.5T systems)
- LC/MS/MS (5 ITMS, 2 QTOF systems)
- DREAMS technology, other innovations

## Data management & data analysis (informatics)

- Proteomics Research Information Storage & Management (PRISM)
- ICR Tools (ICR-2LS)
- Extensive FTICRMS data management experience (15+ years)



Battelle

Philip Morris/TNW/PNNL

6

# Advantages & limitations of newer MS approaches

## Advantages

Higher throughput

More mass accuracy

More sensitivity

Smaller sample requirements

Large dynamic range

## Limitations

Sensitivity < 1 fmole

Relatively expensive

Requires informatic capabilities

Not quantitative

Battelle

Philip M. J. Williams/TNMI/PMN/13

# Need to Remember: Proteomic technology is complex and the limitations are not trivial

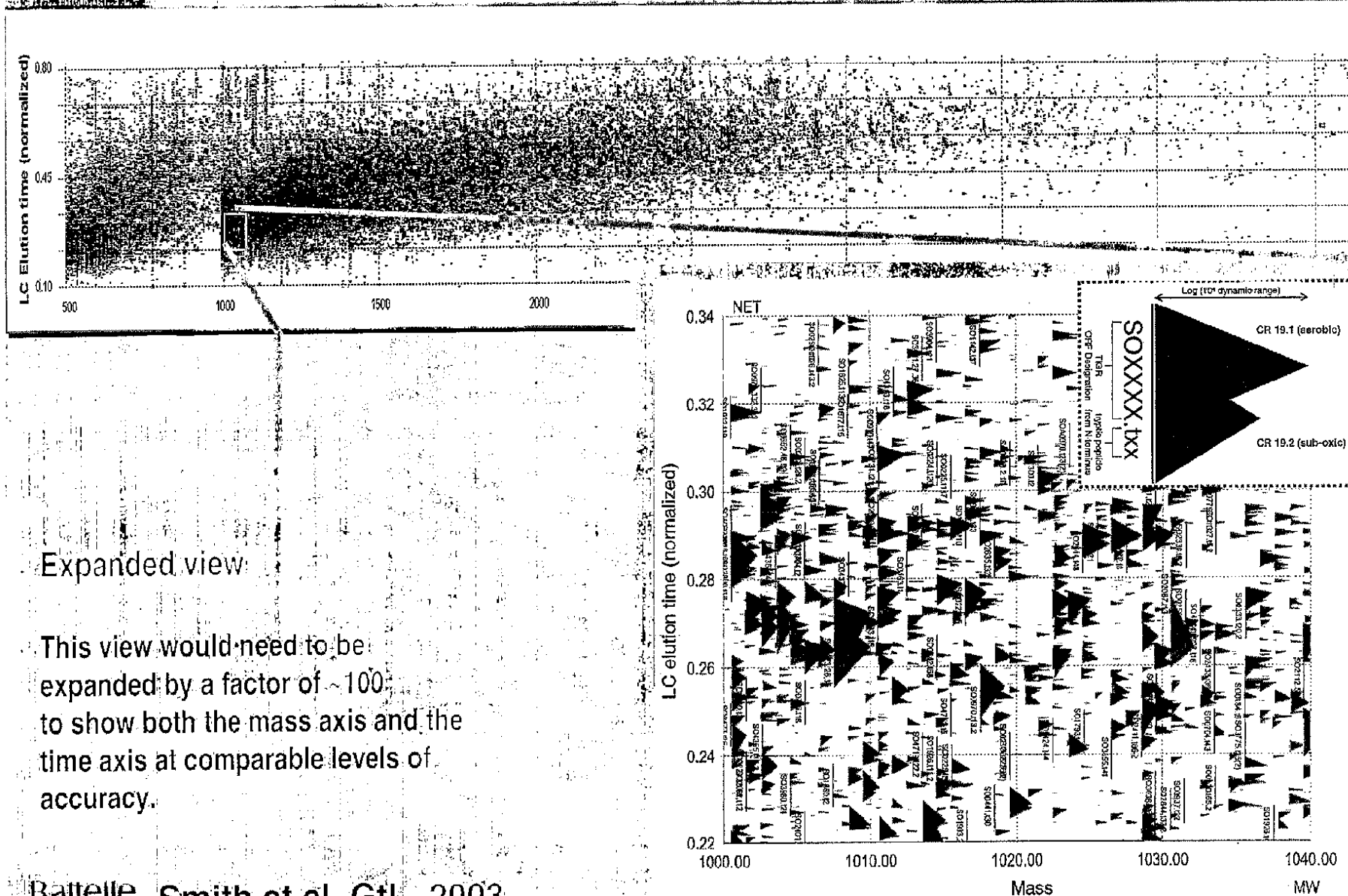
- Chromatographic resolution of proteins, quantitate proteins, identify & characterize proteins
- Post-translational modifications
- Bioinformatics, databases, and data analysis (improved ORF algorithms; functional annotation, etc)
- Sample preparation issues (many)
- Experimental design is critical to success
- Quality control
- Data overload
- "Global proteome" analysis is a myth

Battelle

Philip Morris/TNW/PNNL



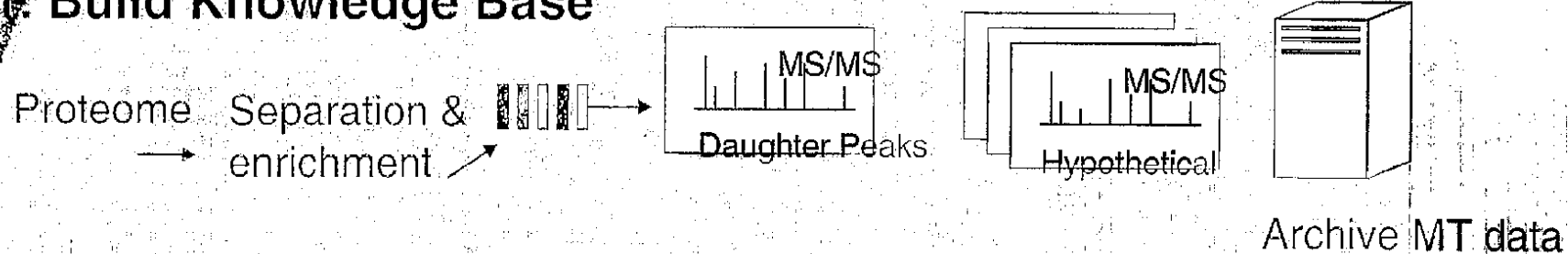
The high mass measurement accuracy and increased sensitivity of FTICR-MS is used to unambiguously identify and/or confirm peptides in a complex mixture



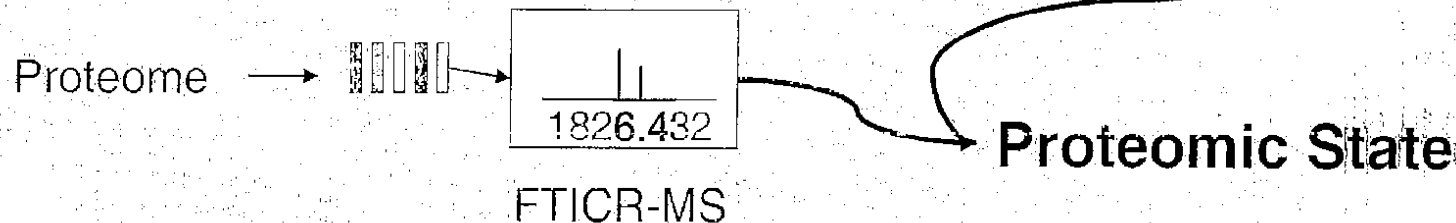
# Proteomics & biomarkers at PNNL

No gels!

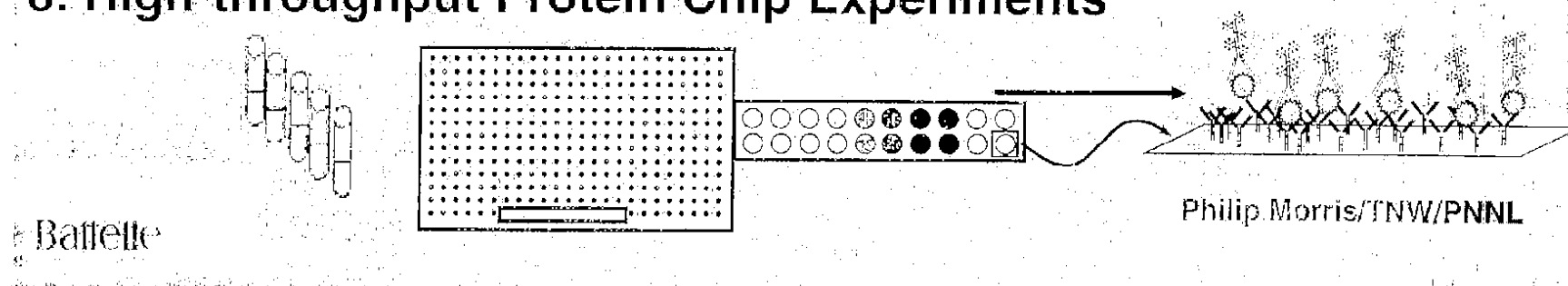
## 1. Build Knowledge Base



## 2. High-throughput FTICR-MS Experiments



## 3. High-throughput Protein Chip Experiments



# Generation and Validation of Accurate Mass & Tags

...APEEKARGITINTAHVEYQTETRHVSHVDCPGHADYVK...

Fragment from Elongation factor Tu

Tryptic digest mixture for analysis

APEEKARG

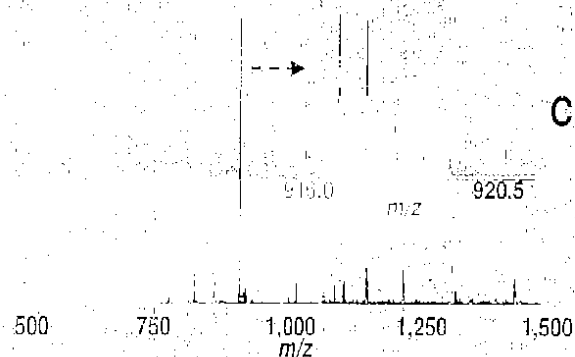
GITINTAHVEYQTETR

RVSHVDCPGHADYVK

Capillary LC-MS/MS selection of peptide

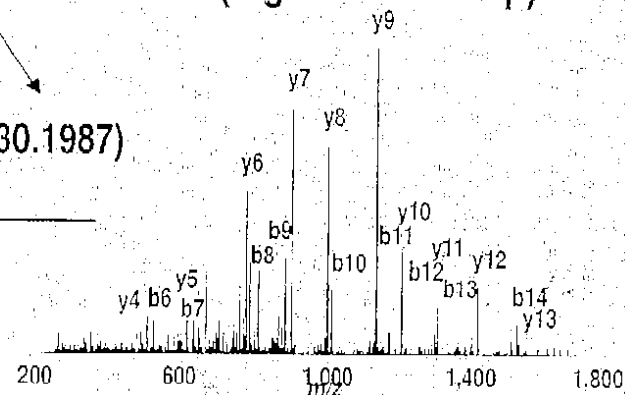
GITINTAHVEYQTETR

Capillary LC- FTICR



Accurate mass observed: 1830.1986

MS/MS (e.g. with ion trap)



Calculated PMT mass (1830.1987)  
and elution time

Potential Mass Tag (PMT) identified:  
GITINTAHVEYQTETR

Validated AMT (GITINTAHVEYQTETR)

# Sample Processing for Dose-Response Study

## Experimental Design:

Male C57 and ICR mice were exposed 0, 75, 250, or 600  $\mu\text{g}$  WTPM/L cigarette smoke for 7 days.

BALF collected at 12 hr post last exposure.

BALF pooled (ten mice/group).

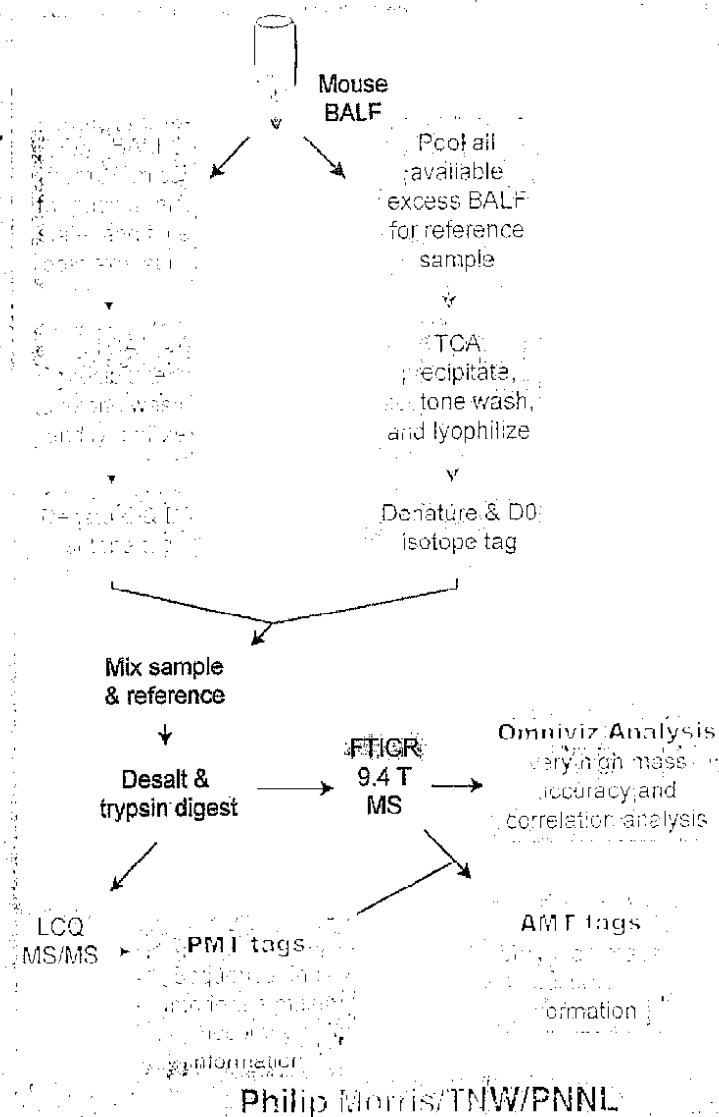
## BALF Sample Preparation:

Pooled to = 40  $\mu\text{g}$  protein,  
precipitated, washed, lyophilized.

## Objectives:

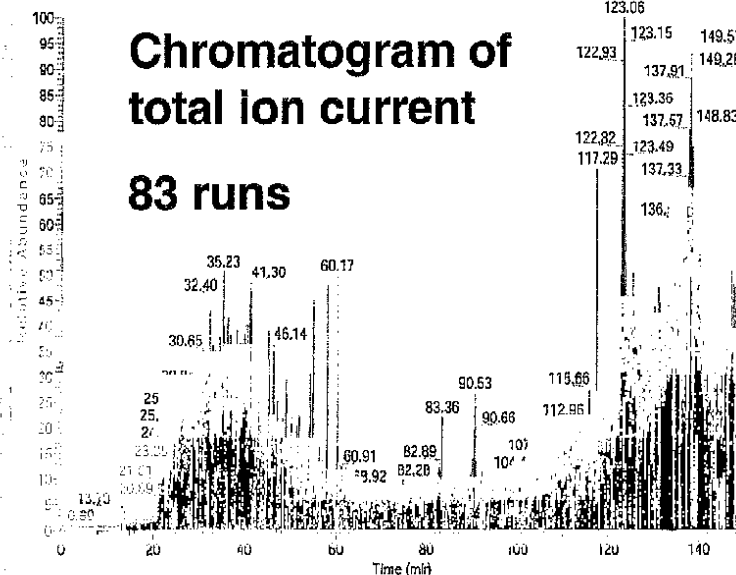
- Adapt methods for BALF sample preparation (not discussed)
- **Characterize protein composition of BALF (proof of technology)**
- **Identify strain- and exposure-dependent changes in BALF protein (proof-of-principle)**

Battelle

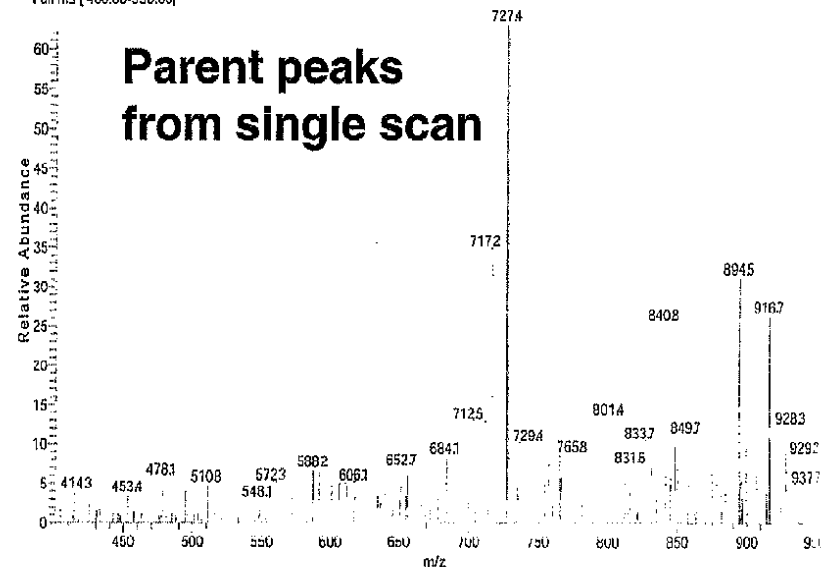


# MS identification of BALF peptide/proteins: LC-MS/MS

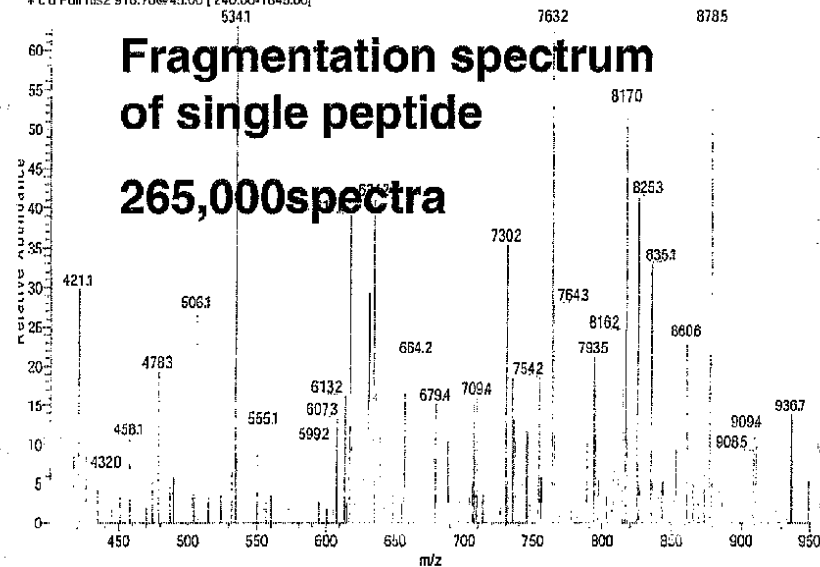
BAL-E1M-12hr\_25oct02\_uranus-a\_0602-4\_400-950.11



BAL-E1M-12hr\_25oct02\_uranus-a\_0602-4\_400-950.11 #1536  
Full ms [400.00-950.00]



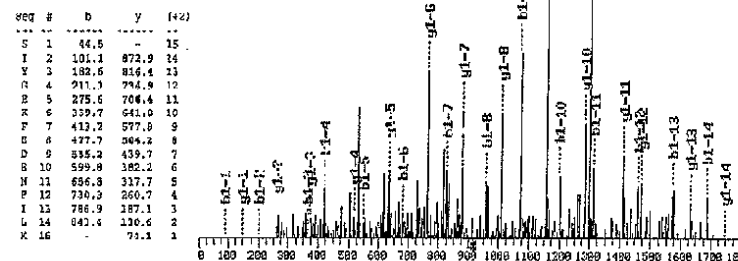
BAL-E1M-12hr\_25oct02\_uranus-a\_0602-4\_400-950.11 #1587  
+ c.d Full ms2 916.70@45.00 [240.00-1645.00]



Seq #	b	y	(+2)
1	89.0	-	15
2	201.1	1744.9	14
3	364.2	1631.8	13
4	421.2	1468.7	12
5	550.2	1411.7	11
6	579.3	1282.6	10
7	625.4	1154.5	9
8	654.5	1087.6	8
9	1059.5	878.4	7
10	1199.0	763.4	6
11	1712.6	634.4	5
12	1488.6	520.3	4
13	1592.7	373.3	3
14	1595.8	260.3	2
15	-	147.1	1

**SEQUEST assignment of spectrum**

**14332 PMT**



## Summary:

### MS identification of peptide/proteins

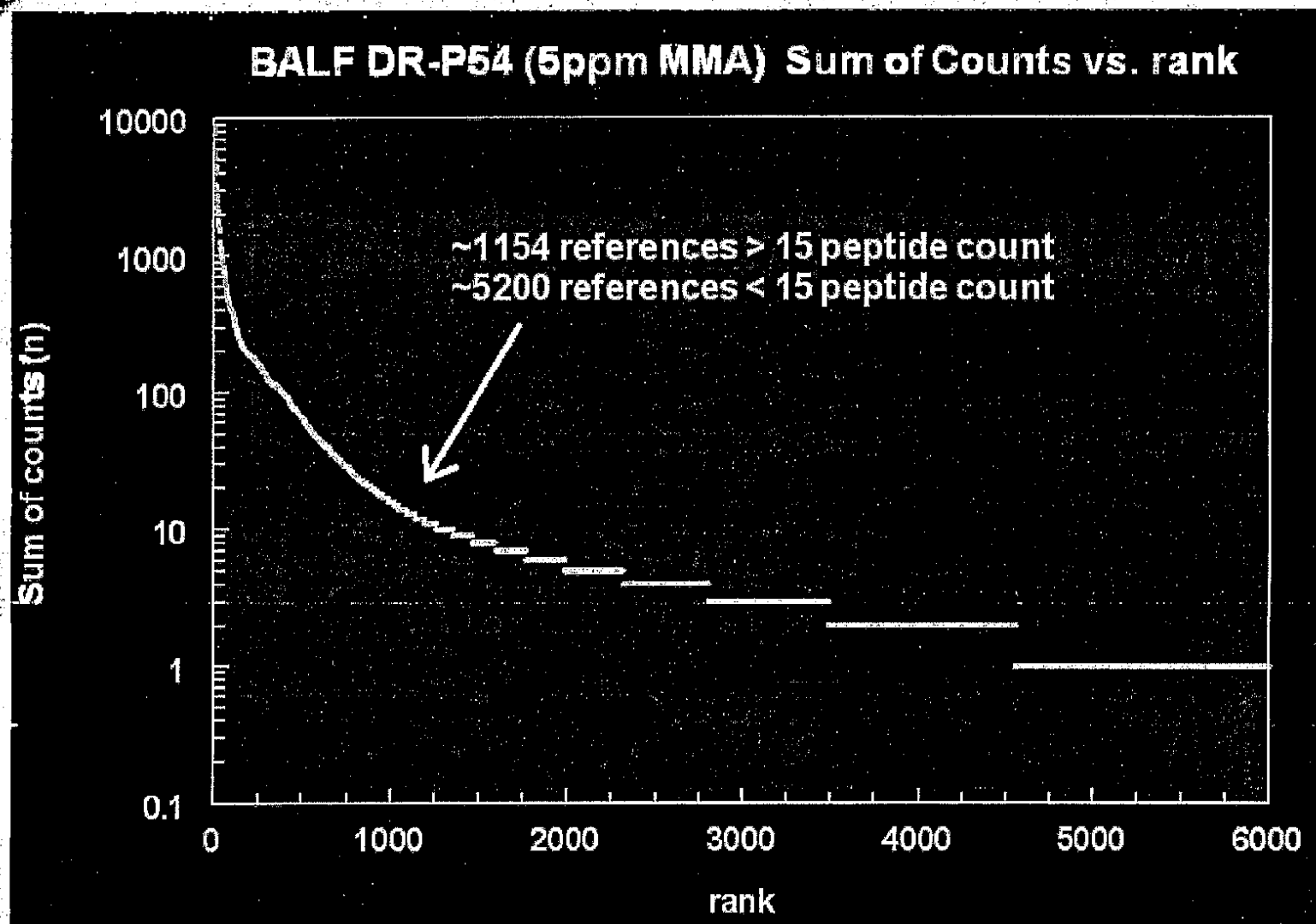
1. To identify 14332 Putative Mass and Time tags (PMT)
2. 27 LC-FTICR-MS runs were analyzed to identify 6284 proteins with at least one AMT and 5461 proteins with AMT that mapped to a single ORF
3. 1365 proteins were identified by 15 or more AMT measurements
4. 1154 / 1365 proteins identified by at least one tryptic AMT in FTICR

Ballell

Adkins et al. (in preparation for Toxicogenomics)

Philip Morris/TNW/PNRL

# Protein identification by peptide count

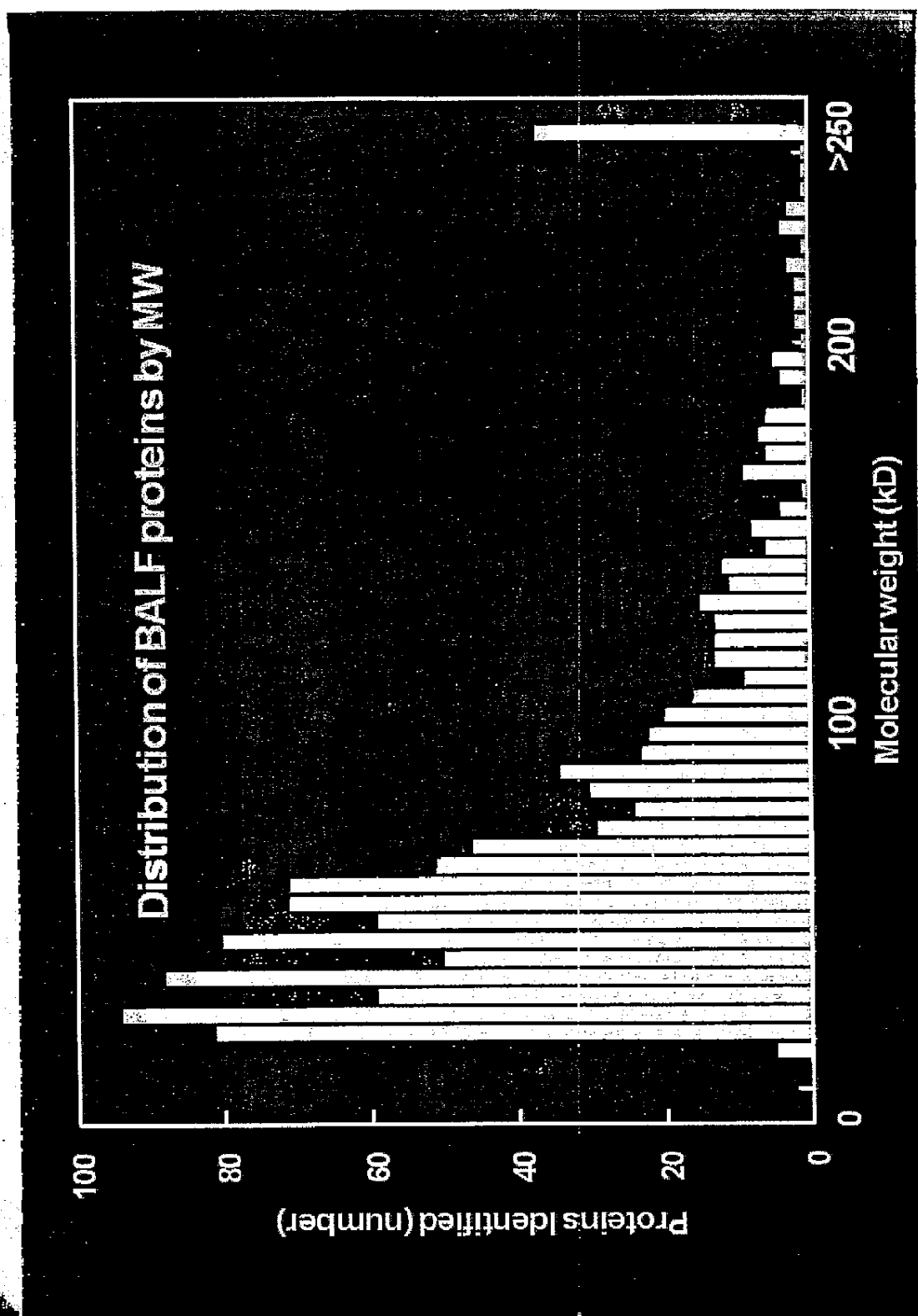


Battelle

Adkins et al. (in preparation for Toxicogenomics)

Philip Morris/TNW/PNNL

PM3001148267



Philip Morris/TNW/PNNL

Adkins et al. (in preparation for *Toxicogenomics*)

Baillie



# Summary of 1154 Identified BALF Proteins

## Most prominent proteins

Serine protease inhibitors 1-1; 1-2, 1-4; 1-4; clade A-3M; clade A-3C

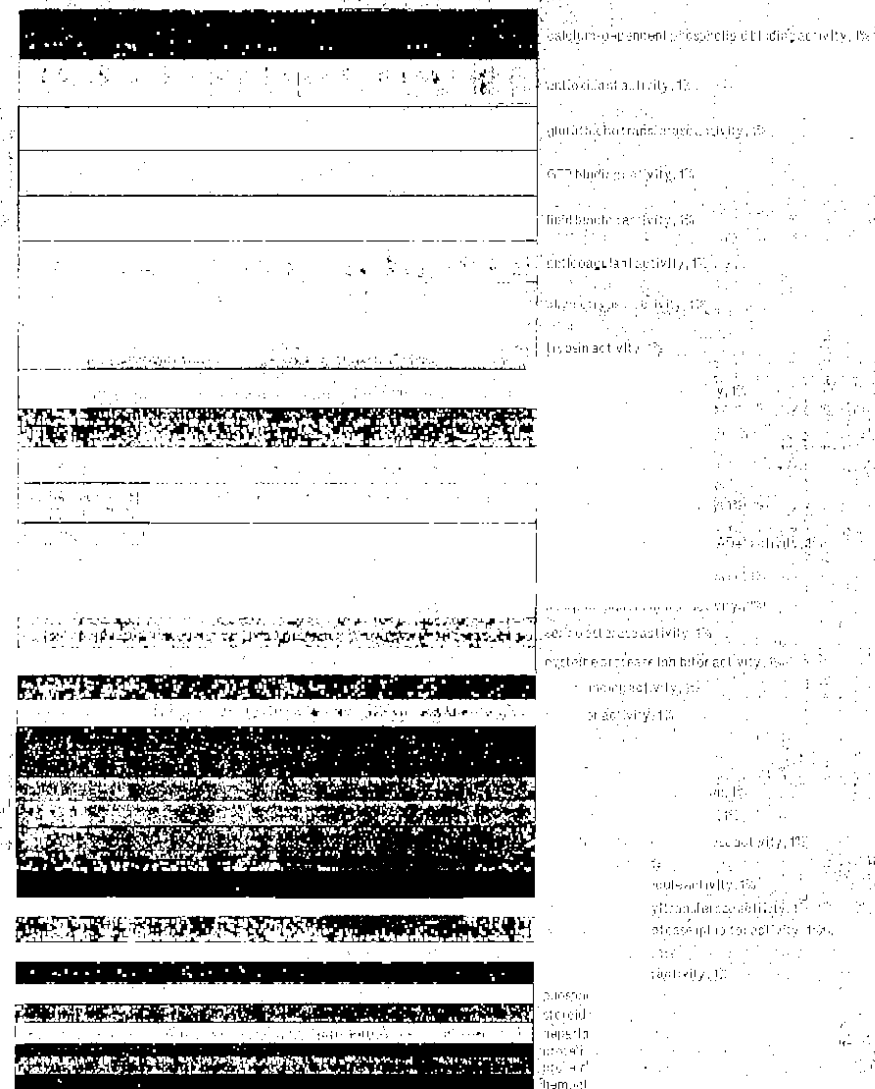
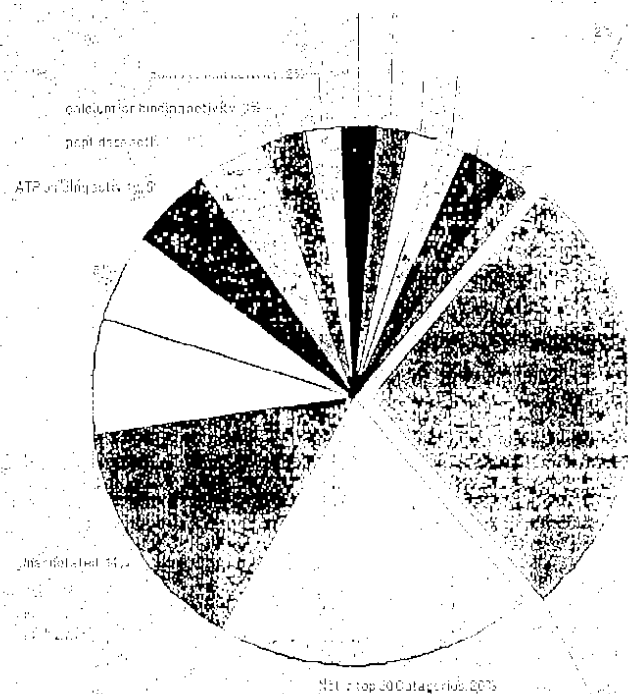
The remaining prominent proteins contain many...

- structural proteins actins or similar to actin
- glycolytic proteins (found in platelets and RBCs)
- typical blood based proteins (e.g. haptoglobin, hemopexin, complement components, transferrin, etc).

Balfelle

Philip Morris/TNW/PNNL

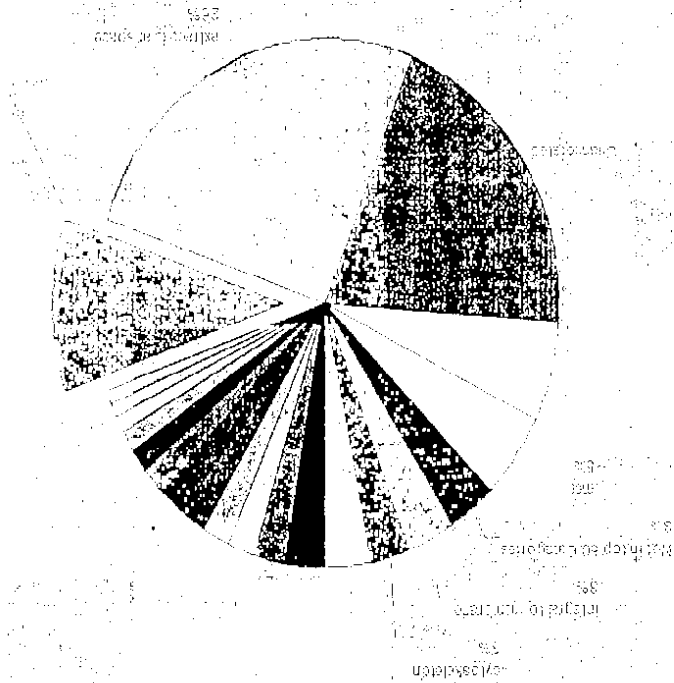
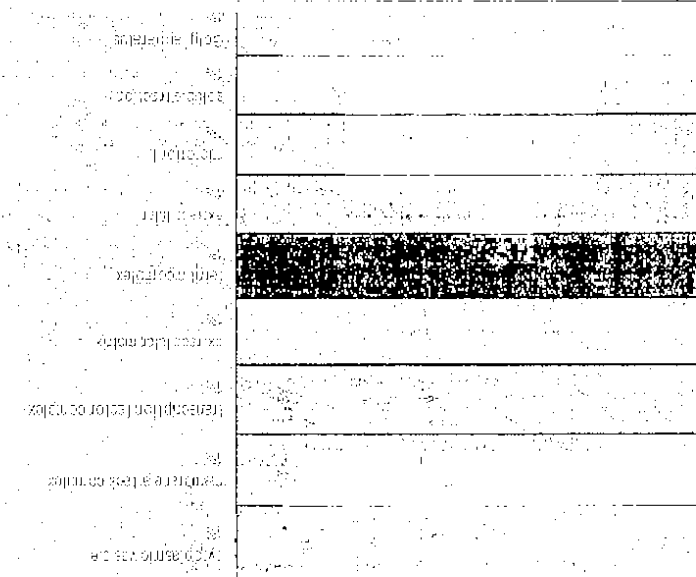
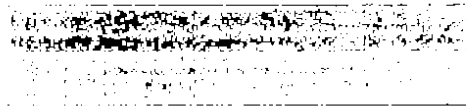
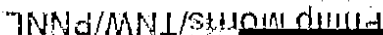
# Gene Ontology Molecular Functions Associated with BALF Proteins



Philip Morris/TNW/PNNL

Battelle

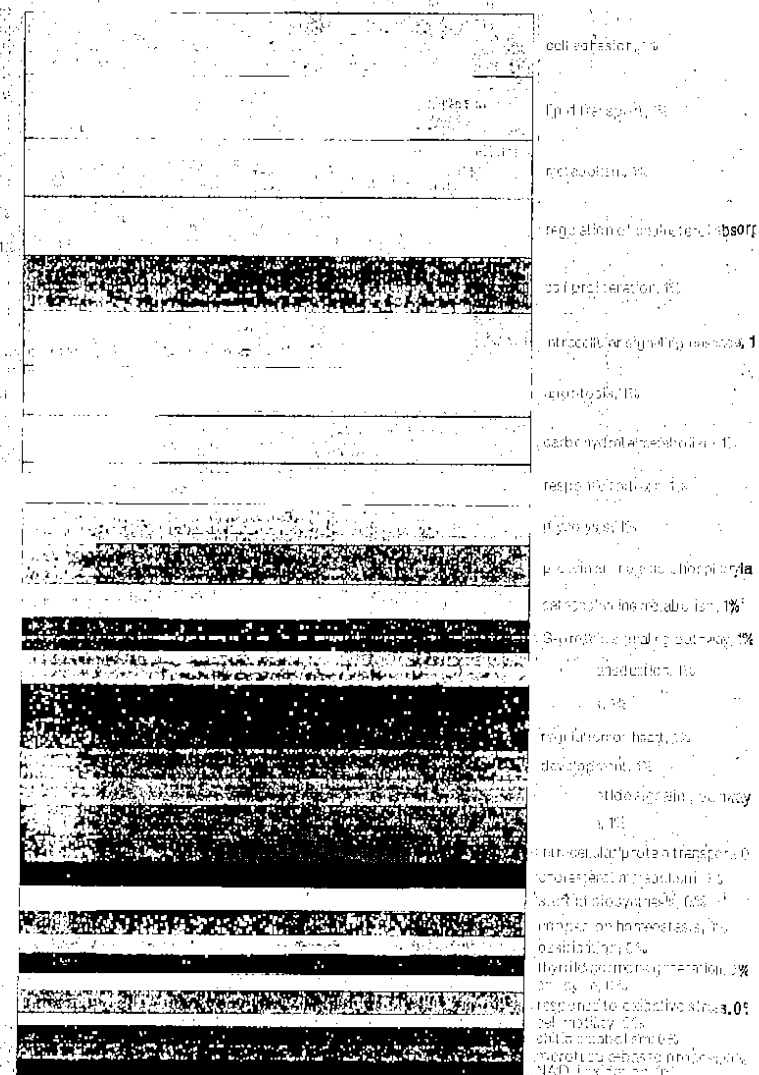
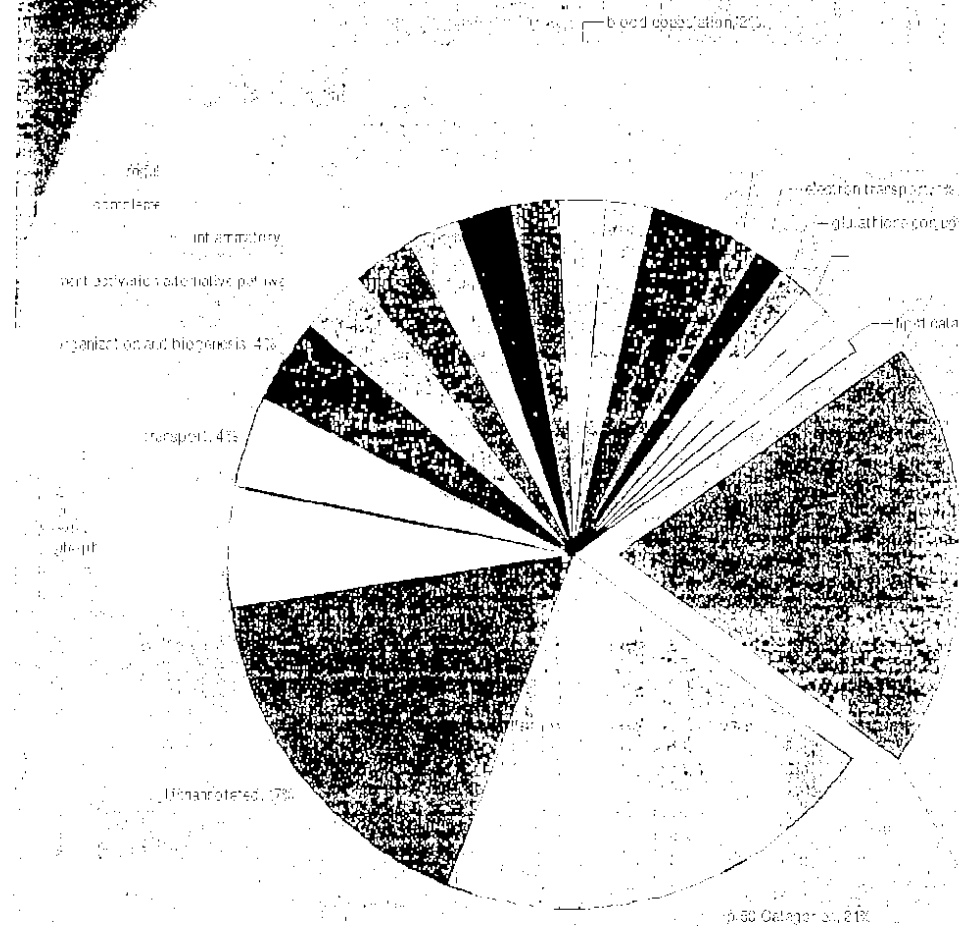
## Ontology Cellular Components



Batterie

PM3001148271

# Gene Ontology Biological Processes Associated with BALF Proteins



Battelle

Philip Morris Inc./PNL

# 53 proteins - related to apoptosis or anti-apoptotic processes

## ► GO Biological Process

31 Apoptosis

13 Anti-apoptosis

9 Induction of apoptosis

## ► GO Molecular Function

9 apoptosis regulator activity

7 apoptosis inhibitor activity

Bellelle

Philip Morris/TNW/PNML

# Apoptosis related proteins

many proteins with well defined roles, others less clear

Remember – all not detected with equal confidence

Battelle

Protein Name	GO Biological Process	GO Molecular Function	GO Cellular Component
angiotensin II receptor; type 2; AT2 receptor	apoptosis	protein kinase inhibitor activity	
apoptosis inhibitor 3	anti-apoptosis	apoptosis inhibitor activity	cytosol
apoptosis inhibitor 5	anti-apoptosis	apoptosis inhibitor activity	spliceosome complex
baculovirus inhibitor of apoptosis repeat containing ubiquitin cycle	ubiquitin cycle	ubiquitin conjugating enzyme activity	
Bcl-2 binding component 3; BH3-only	induction of apoptosis	apoptosis activator activity	mitochondrial membrane
Bcl2-associated athanogene 1	apoptosis	apoptosis regulator activity	cytosol
beta-amyloid binding protein precursor	induction of apoptosis	G-protein coupled receptor activity	integral to plasma membrane
breast cancer 2	apoptosis	single-stranded DNA binding activity	synaptonemal complex
Bruton agammaglobulinemia tyrosine kinase; Bcr	induction of apoptosis	ATP binding activity	cytoplasm
caspase 4: apoptosis-related cysteine protease; proteolysis and peptidolysis	caspase activity		intracellular
caspase 8 associated protein 2; DNA segment: C	induction of apoptosis	ATP binding activity	extracellular space
caspase recruitment domain family; member 10; activation of NF-kappaB	apoptosis	apoptosis regulator activity	kinesin complex
caspase recruitment domain family; member 14	apoptosis	apoptosis regulator activity	kinesin complex
CCAAT/enhancer binding protein (C/EBP); beta; induction of apoptosis	apoptosis	apoptosis activator activity	nucleus
chromosome segregation 1-like (S. cerevisiae)	apoptosis	importin-alpha export receptor activity	cytoplasm
complement component 9	apoptosis	complement activity	extracellular space
cullin 3	induction of apoptosis		
death associated protein kinase 1	induction of apoptosis	ATP binding activity	cytoplasm
death effector domain-containing; tumor necrosis	induction of apoptosis	DNA binding activity	cytoplasm
death inducer-obliterator 1	apoptosis	DNA binding activity	cytoplasm
deleted in colorectal carcinoma	induction of apoptosis		integral to plasma membrane
deoxyribonuclease II alpha; deoxyribonuclease I	apoptosis		lysosome
engulfment and cell motility 2; ced-12 homolog (C. elegans)	apoptosis		
eukaryotic translation elongation factor 1 alpha	anti-apoptosis		cellular component unknown
expressed sequence A1324246	apoptosis	ATP binding activity	eukaryotic translation elongation
Fas-activated serine/threonine kinase	apoptosis	RNA binding activity	
glial cell line derived neurotrophic factor; glial cell	anti-apoptosis	protein serine/threonine kinase activity	cytoplasm
glycogen synthase kinase 3 beta	anti-apoptosis		
granzyme G; AKA granzyme G; CTL serine protease	apoptosis	ATP binding activity	cytoplasm
helicase; lymphoid specific; proliferation-associated	anti-apoptosis	DNA helicase activity	
hemolytic complement	apoptosis	complement activity	extracellular space
lectin; galactose binding; soluble 1	apoptosis	galactose binding activity	cellular component unknown
lymphotoxin B receptor; LT beta-R; LT-beta receptor	apoptosis	receptor activity	extracellular space
mast cell maturation inducible protein 1	apoptosis		
neuronal apoptosis inhibitory protein; related serine	anti-apoptosis	apoptosis inhibitor activity	integral to membrane
nuclear factor of kappa light chain gene enhancer	apoptosis	transcription factor activity	cytoplasm
phosphodiesterase 1B; Ca2+-calmodulin dependent	apoptosis	calmodulin binding activity	insoluble fraction
phosphoprotein enriched in astrocytes 15; mammalian	apoptosis	apoptosis regulator activity	cytosol
pleiomorphic adenoma gene-like 1	apoptosis	nucleic acid binding activity	transcription factor complex
programmed cell death 2		apoptosis regulator activity	nucleus
programmed cell death 6 interacting protein	apoptosis	signal transducer activity	cytosol
protein kinase C: epsilon	induction of apoptosis	calcium independent protein kinase activity	
protein kinase raf 1; murine sarcoma 3611 oncogene	apoptosis	ATP binding activity	cytosol
RIKEN cDNA 8330415L08	apoptosis	apoptosis regulator activity	
RIKEN cDNA B430311C09	apoptosis	apoptosis regulator activity	nucleus
secreted frizzled-related sequence protein 1	anti-apoptosis	transmembrane receptor activity	extracellular space
seven in absentia 1A	apoptosis	transcription co-repressor activity	cytoplasm
seven in absentia 1B	apoptosis	transcription co-repressor activity	cytoplasm
SH3-domain GRB2-like B1 (endophilin)	apoptotic program	apoptosis activator activity	cytosol
similar to interleukin 19	induction of apoptosis		
TGF-beta1-induced anti-apoptotic factor 1; myosin	apoptosis	ATP binding activity	myosin
transformation related protein 63; KET protein	induction of apoptosis		
tumor necrosis factor receptor superfamily; member 1	anti-apoptosis	receptor activity	extracellular space
tumor protein; translationally-controlled 1; translationally	anti-apoptosis	apoptosis inhibitor activity	cellular component unknown
UNC-13 homolog (C. elegans) 1	induction of apoptosis		cellular component unknown

# Selected proteins related to immune response

Remember – all not detected with equal confidence

Battelle

Protein/Notes	GO Biological Process 1	GO Molecular Function 1	GO Cellular Component 1
5' oligoadenylate synthetase 1A	immune response	RNA binding activity	
5' oligoadenylate synthetase-like 1; 2:5-oligoadenylate synthetase-like 1	immune response	ATP binding activity	
allograft inflammatory factor 1		molecular function unknown	centriole
alpha-2-glycoprotein 1; zinc	immune response	zinc ion binding activity	extracellular space
atractin	inflammatory response	structural molecule activity	integral to plasma membrane
basic helix-loop-helix domain containing: class B2 negative regulation of transcription	negative regulation of transcription	DNA binding activity	nucleus
Bruton agammaglobulinemia tyrosine kinase; Bruton's tyrosine kinase	intracellular signaling cascade	protein serine/threonine kinase activity	
Burkitt lymphoma receptor 1	G-protein coupled receptor	C-X-C chemokine receptor activity	integral to plasma membrane
cardiotrophin 1	immune response	cytokine activity	extracellular
D83 antigen	humoral immune response		extracellular space
chemokine (C-C) receptor 3; chemokine (C-C) receptor 3			extracellular space
chemokine (C-X-C) KC; GRO	immune response		extracellular space
chemokine (C-X-C motif) ligand 15; small inducible chemokine	signal transduction		extracellular space
chemokine (C-X-C motif) ligand 15; small inducible chemokine	hemopoiesis	chemokine activity	extracellular space
colony stimulating factor 1; colony-stimulating factor 1	positive regulation of cell proliferation	cytokine activity	extracellular space
colony stimulating factor 2 receptor: beta 2; low-affinity receptor	signal transduction	hematopoietin/interferon-class III cytokine receptor activity	extracellular space
complement component 3; complement factor 3; a1-inhibitor of complement	inflammatory response	endopeptidase inhibitor activity	
conserved helix-loop-helix ubiquitous kinase; IKK	immune response	IkappaB kinase activity	
cornichon homolog	immune response	calcium ion binding activity	integral to membrane
growth differentiation factor 3	cell growth and/or maintenance	cytokine activity	extracellular space
growth hormone receptor		hematopoietin/interferon-class III cytokine receptor activity	extracellular space
histocompatibility 2: M region locus 10.1	immune response	defense/immunity protein activity	extracellular space
immune associated nucleotide 4	mRNA splicing	ATP binding activity	integral to membrane
interferon activated gene 202A; Interferon activated gene 202A	immune response	DNA binding activity	nucleus
interferon gamma receptor; INF-g receptor	immune response	hematopoietin/interferon-class III cytokine receptor activity	extracellular space
interferon-induced protein with tetratricopeptide repeat	immune response	transferase activity	cytoplasm
interleukin 1 family: member 6; Interleukin 1 family: member 6	inflammatory response	interleukin-1 receptor ligand activity	cellular component unknown
interleukin 1 receptor antagonist	immune response		extracellular
interleukin 11 receptor: alpha chain 2; locus 2			integral to plasma membrane
interleukin 18 binding protein; Interferon gamma 1	immune response	interleukin-18 binding activity	cellular component unknown
interleukin 18 binding protein; Interferon gamma 1			extracellular space
interleukin 5 receptor: alpha; IL-5 receptor alpha (cell surface receptor linked to electron transporter activity)		electron transporter activity	extracellular space
interleukin-1 receptor-associated kinase 3			
kininogen; H-kininogen; L-kininogen	inflammatory response		
leukotriene A4 hydrolase; LTA4 hydrolase	inflammatory response	epoxide hydrolase activity	integral to plasma membrane
lymphocyte-activation gene 3	defense response	interleukin-1, Type II, blocking activity	extracellular space
lysozyme	humoral immune response		
macrophage activation 2	immune response	ATP binding activity	cytosol
odd Oz/ten-m homolog 1	immune response	heparin binding activity	extracellular
protein-tyrosine sulfotransferase 1	humoral immune response		
protein-tyrosine sulfotransferase 1	immune response	ATP binding activity	integral to membrane
proteoglycan 3; major basic protein 2	inflammatory response	protein-tyrosine sulfotransferase activity	Golgi apparatus
resistin like alpha; found in inflammatory zone 1	immune response	extracellular matrix structural	extracellular space
S100 calcium binding protein A9 (calgranulin B); S100 calcium binding protein A9	cell proliferation	hormone activity	extracellular space
signal transducer and activator of transcription 4	inflammatory response	calcium ion binding activity	extracellular space
signal transducer and activator of transcription 5; IFN-gamma inducible	immune response	enzyme activity	
signal transducer and activator of transcription 5; IFN-gamma inducible	immune response	translation regulator activity	
signal transducer and activator of transcription 5; IFN-gamma inducible	immune response	signal transducer activity	
similar to allograft inflammatory factor 1			
similar to Chain L: Antigen-Antibody Complex	humoral immune response	antigen binding activity	
solute carrier family 11 (proton-coupled divalent cation symporter)	immune response	iron ion transporter activity	integral to membrane
T cell cytokine receptor; cytokine receptor family: immune response	immune response	MHC protein binding activity	extracellular space
T cell cytokine receptor; cytokine receptor family:		interleukin-27 receptor activity	
tol-like receptor 1	macrophage activation	interleukin-1, Type II, activating activity	phagocytosis
tol-like receptor 7	inflammatory response	interleukin-1, Type II, activating activity	extracellular space
tol-like receptor 9	inflammatory response	transmembrane receptor activity	extracellular space
transporter 1: ATP-binding cassette: sub-family B; immune response	immune response	ATP-binding cassette (ABC) transporter activity	extracellular space
tumor necrosis factor (ligand) superfamily: member 1	immune response	tumor necrosis factor receptor activity	integral to membrane

# Selected proteins related to acute phase response

Remember - not all not detected with equal confidence

FirstOfNotes	GO Molecular Function 1	GO Cellular Component 1	GO Biological Process 1
coagulation factor II; prothrombin	calcium ion binding activity	extracellular space	acute-phase response
coagulation factor VIII; Factor VIII	ATP binding activity	extracellular space	acute-phase response
haptoglobin	chymotrypsin activity	extracellular matrix	acute-phase response
hemopexin		extracellular space	acute-phase response
interleukin 6	interleukin-6 receptor ligand	extracellular space	acute-phase response
orosomucoid 1	transporter activity	extracellular space	acute-phase response
orosomucoid 2	transporter activity	extracellular space	acute-phase response
serine protease inhibitor 1-2; DNA seg	peptidase activity	extracellular space	acute-phase response
serine protease inhibitor 1-4	peptidase activity	extracellular space	acute-phase response
serine protease inhibitor 1-5	peptidase activity	extracellular space	acute-phase response
serum amyloid A 1	acute-phase response protei	extracellular	acute-phase response

Battelle

Philip Morris/TNW/PNNL

24



Many of the identified proteins are involved in proteolysis or peptidolysis

~70 of 120 shown here

- numerous proteins from RIKEN or EST not shown

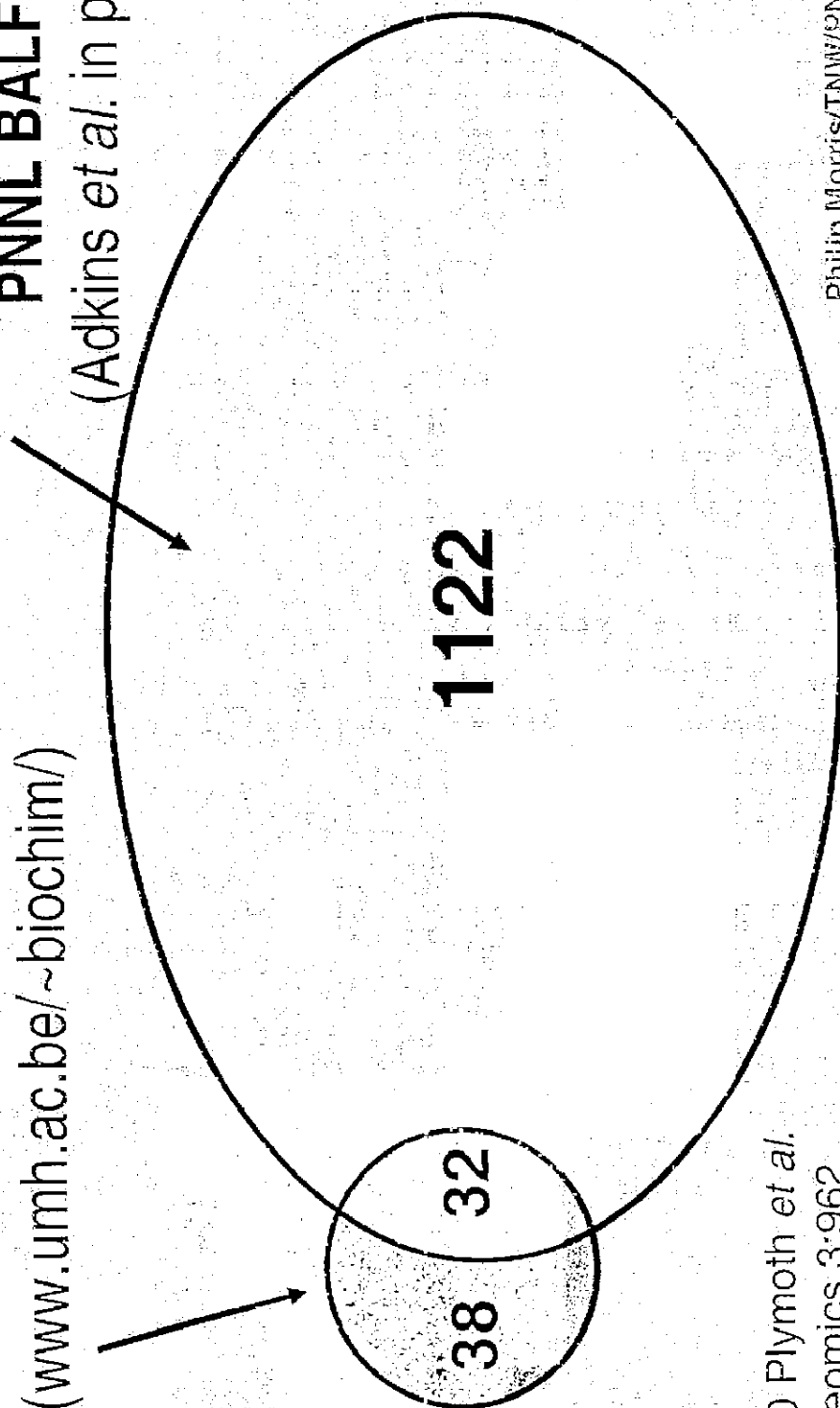
Battelle

First Of Notes	GO Molecular Function	GO Cellular Component	GO Biological Process
disintegrin and metalloprotease domain 10; a	metalloendopeptidase activity	extracellular space	proteolysis and peptidolysis
disintegrin and metalloprotease domain 11; a	metalloendopeptidase activity	integral to membrane	proteolysis and peptidolysis
disintegrin and metalloprotease domain 2 (fer)		integral to membrane	proteolysis and peptidolysis
disintegrin and metalloprotease domain 23	metalloendopeptidase activity		proteolysis and peptidolysis
disintegrin and metalloprotease domain 25 (te)	integrin binding activity	extracellular matrix	proteolysis and peptidolysis
disintegrin and metalloprotease domain 26 (te)	metalloendopeptidase activity	integral to membrane	proteolysis and peptidolysis
disintegrin and metalloprotease domain 33	zinc ion binding activity	extracellular space	proteolysis and peptidolysis
disintegrin-like and metalloprotease (repolys)	metalloendopeptidase activity	plasma membrane	proteolysis and peptidolysis
adipsin; D component (adipsin) of complement	complement factor D activity	extracellular space	proteolysis and peptidolysis
aminocyclase 1	metallopeptidase activity		proteolysis and peptidolysis
arginyl aminopeptidase (aminopeptidase B); sim	aminopeptidase B activity		proteolysis and peptidolysis
calpain 10; calpain 8	calpain activity	intracellular	proteolysis and peptidolysis
calpain 6	calpain activity	cytoplasm	proteolysis and peptidolysis
calpain 7	calpain activity	cytoplasm	proteolysis and peptidolysis
calpain 8; stomach-specific calpain (nCL-2)	calcium ion binding activity	intracellular	proteolysis and peptidolysis
calpastatin	calpain inhibitor activity	cellular component unk	protein catabolism
caspase 4; apoptosis-related cysteine protease;	caspase activity	intracellular	proteolysis and peptidolysis
cathepsin 3 precursor	molecular function unknown	lysosome	proteolysis and peptidolysis
cathepsin B preproprotein	peroxidase activity	lysosome	proteolysis and peptidolysis
cathepsin D	cathepsin D activity	extracellular space	proteolysis and peptidolysis
cathepsin G preproprotein	cathepsin G activity	extracellular space	proteolysis and peptidolysis
cathepsin H; Cat H	cathepsin H activity	extracellular space	proteolysis and peptidolysis
cathepsin S preproprotein; Cat S	cathepsin S activity	extracellular space	proteolysis and peptidolysis
cathepsin Z preproprotein; cathepsin Z precurs	cysteine-type endopeptidase acti	extracellular space	proteolysis and peptidolysis
cytoplasmic nuclear factor of activated T-cells 3	cysteine-type endopeptidase acti	nucleus	proteolysis and peptidolysis
endothelin converting enzyme 2	neprilysin activity	nucleosome	proteolysis and peptidolysis
granzyme G; AKA granzyme G; CTL serine prote			proteolysis and peptidolysis
granzyme K	chymotrypsin activity	cytoplasm	proteolysis and peptidolysis
granzyme M (lymphocyte met-ase 1); lymphocyte	trypsin activity	extracellular space	proteolysis and peptidolysis
heptoglobin	hemoglobin binding activity	extracellular space	proteolysis and peptidolysis
hepsin	chymotrypsin activity	integral to membrane	proteolysis and peptidolysis
histocompatibility 2; complement component fac	chymotrypsin activity		proteolysis and peptidolysis
kallikrein 5	chymotrypsin activity	extracellular space	proteolysis and peptidolysis
leukotriene A4 hydrolase; LTA4 hydrolase	metalloendopeptidase activity		proteolysis and peptidolysis
low density lipoprotein receptor-related protein	trypsin activity		proteolysis and peptidolysis
mast cell protease 1	chymotrypsin activity	cytoplasm	proteolysis and peptidolysis
mast cell protease 2	chymotrypsin activity	cytoplasm	proteolysis and peptidolysis
membrane-bound transcription factor protease;		endoplasmic reticulum	proteolysis and peptidolysis
nepril 1 beta; nepril beta	astacin activity	extracellular space	proteolysis and peptidolysis
methionine aminopeptidase 2; eIF-2-associated			proteolysis and peptidolysis
methionyl aminopeptidase 1	methionyl aminopeptidase activit		proteolysis and peptidolysis
monoglyceride lipase; EST AA589436			proteolysis and peptidolysis
myosin VIb	motor activity		proteolysis and peptidolysis
nuclear receptor interacting protein 2	protein binding activity		proteolysis and peptidolysis
odd Oz/ten-m homolog 1			proteolysis and peptidolysis
O-sialoglycoprotein endopeptidase	O-sialoglycoprotein endopeptidase	cytoplasm	proteolysis and peptidolysis
aplin; proteoglycan-like sulfated glycoprotein	enzyme activity	basement membrane	proteolysis and peptidolysis
phosphate regulating gene with homologies to	zinc ion binding activity		proteolysis and peptidolysis
phospholipase A2; group IVA (cytosolic; calcium	phospholipase A2 activity		proteolysis and peptidolysis
plasminogen activator; urokinase; urokinase-ty	U-plasminogen activator activity	extracellular space	proteolysis and peptidolysis
procollagen: type V: alpha 1; pro-alpha1(V) colla	heparin binding activity	extracellular space	proteolysis and peptidolysis
proprotein convertase subtilisin/kexin type 2; p	proprotein convertase 2 activity	extracellular space	proteolysis and peptidolysis
proprotein convertase subtilisin/kexin type 5	subtilase activity	integral to membrane	proteolysis and peptidolysis
proprotein convertase subtilisin/kexin type 7	ATP binding activity	extracellular space	proteolysis and peptidolysis
protease: serine: 11 (Igf binding); insulin-like	growth factor binding		proteolysis and peptidolysis
protease: serine: 12 neurotrypsin; (motopsin)	chymotrypsin activity	extracellular space	proteolysis and peptidolysis
protease: serine: 14 (epithlin)	tRNA ligase activity	integral to plasma memb	proteolysis and peptidolysis
protease: serine: 18	chymotrypsin activity	extracellular space	proteolysis and peptidolysis
protease: serine: 7 (enterokinase); enterokinase	chymotrypsin activity	integral to membrane	proteolysis and peptidolysis
proteasome 26S non-ATPase subunit 8		proteasome regulatory p	proteolysis and peptidolysis
psin 2 tandem duplication of Ren1	pepsin A activity		proteolysis and peptidolysis
serine (or cysteine) proteinase inhibitor; clade	Eserine protease inhibitor activity	cytoplasm	regulation of proteolysis &
serine protease inhibitor 12	serine protease inhibitor activity	cytosol	regulation of proteolysis &
suppressor of Ty 16 homolog; suppressor of Ty	metalloexopeptidase activity		proteolysis and peptidolysis
thyrotropin-releasing hormone degrading ectoe	membrane alanyl aminopeptidase		proteolysis and peptidolysis
Tnf receptor-associated factor 6	nucleic acid binding activity		proteolysis and peptidolysis
tolloid-like	astacin activity	extracellular space	proteolysis and peptidolysis

# MS identification of peptide/proteins (database comparisons)

Wattiez et al. – human BALF  
([www.umh.ac.be/~biochim/](http://www.umh.ac.be/~biochim/))

Philip Morris / TNW /  
PNNL BALF  
(Adkins *et al.* in prep)



>100 Plymoth *et al.*  
Proteomics 3:962,  
2003

Philip Morris/TNW/PNNL

# Can MS proteomics identify mouse strain- and exposure-dependent changes in BALF?

BALF pooled within treatment group

FTICR analyses in triplicate

OmniVizPro – exploratory data analysis and  
visualization using K-means (not statistics!!!)

Baillie

Philip Morris/INW/PNUL

# Data Transformation

Analysis	Control		Low		Medium		High	
All data	C57	ICR	C57	ICR	C57	ICR	C57	ICR
C57 vs. ICR	C57	ICR					C57	ICR
C57	C57		C57		C57		C57	
ICR		ICR		ICR		ICR		ICR

## 1. Log(10) abundance (current per peptide)

	C57-0	C57-75	C57-250	C57-600
Peptide	5.7e7	5.9e7	6.2e7	6.1e7

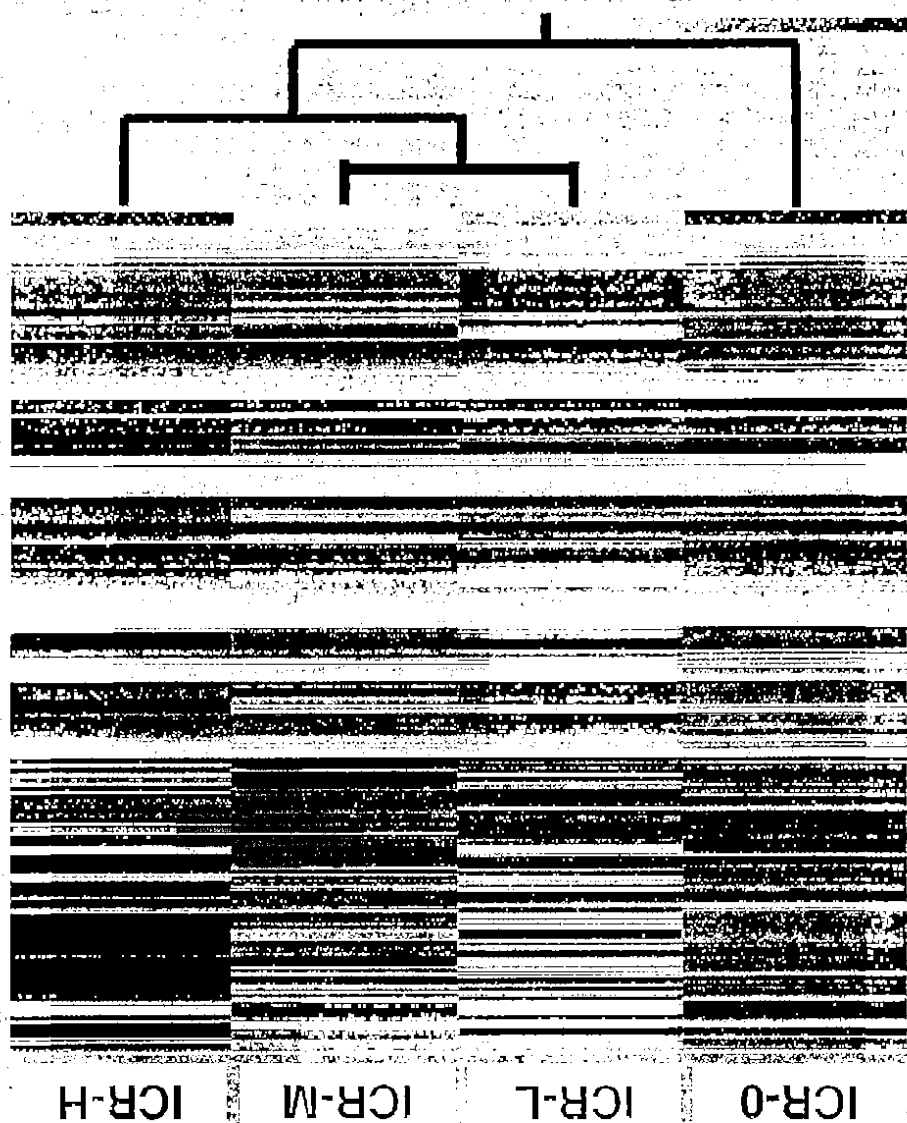
## 2. Normalize abundance to mean value of row.

Peptide	0.92	1.05	1.23	1.15
---------	------	------	------	------

Battelle

Philip Morris/TNW/PMNL

Can we see a dose difference with in ICR mice?



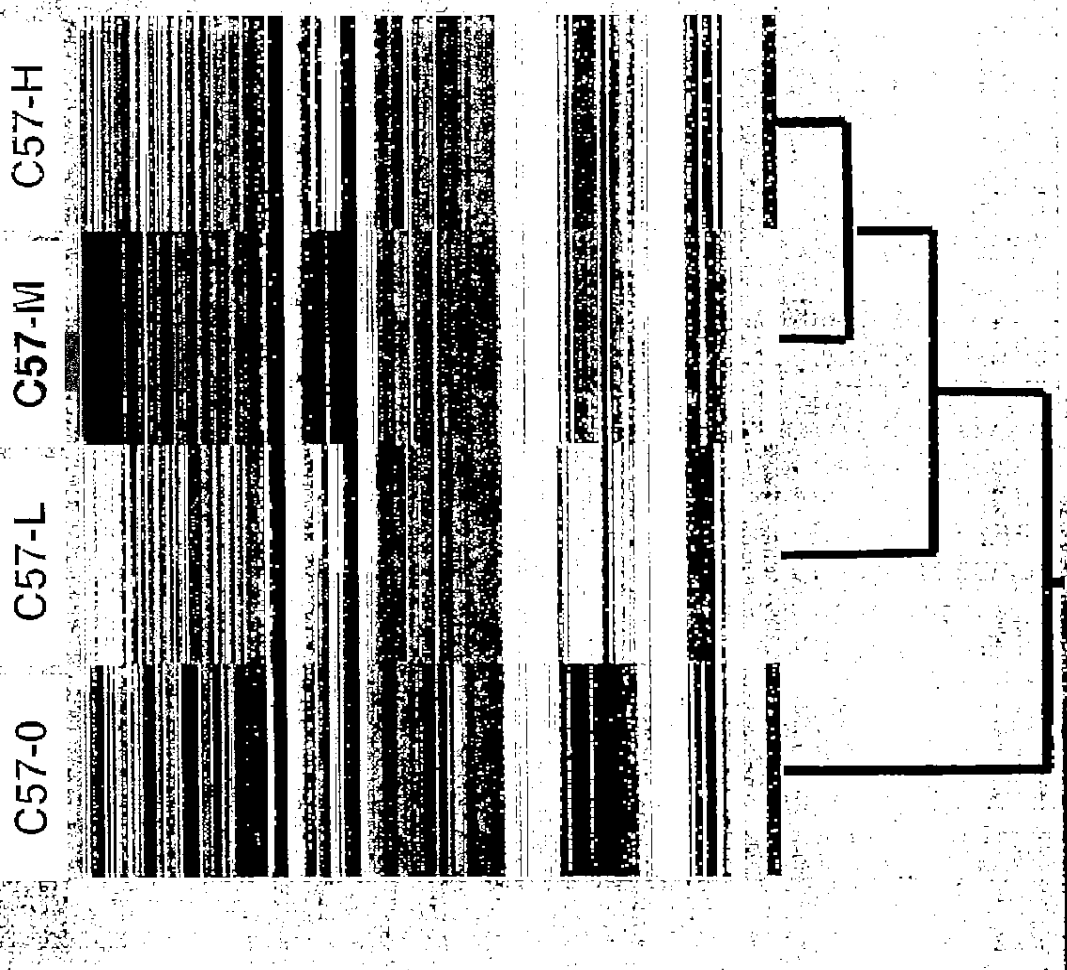
Yes, high level  
dendrogram of  
four groups sorts  
them to logical  
dose-response  
order.

K-means, Log10

Battelle

Can we see a dose difference with in C57 mice?

[chemrxiv.org/content/2018/04/01/10.26434/chemrxiv-2018-04-01-10.26434](http://chemrxiv.org/content/2018/04/01/10.26434/chemrxiv-2018-04-01-10.26434)



Yes, high level  
dendrogram of  
four groups sorts  
them to logical  
dose-response  
order

K-means, Log10

Bathelle

# KC and TARC

**KC** (Growth Related Oncogene-alpha, GRO<sub>mouse</sub>)

Chemokine with C-X-C motif

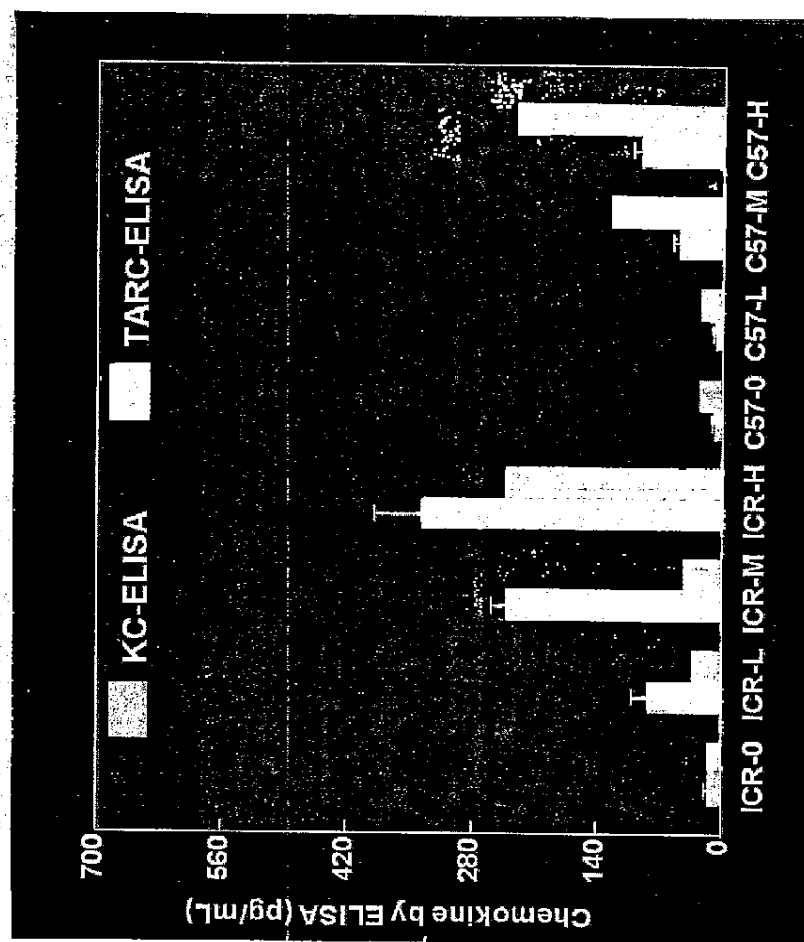
Inducible by PDGF, secretory protein,

1 mipatrsflc aalllllatsr latgapiane lrcc **clqtma gihll** niosl  
 xvlpsghpct qteviatlkn greacldpea plvgkiqkm lkgvpk

Theoretical pI: 8.23 / Mw  
 (average mass): 1214.51

This nontryptic peptide had a decent Xcorr (best was 2.57) and was also seen in forty individual LCQ scans.

Batlle



# ELISA vs. FTICR-MS ?

Note: Ion current  
scaled

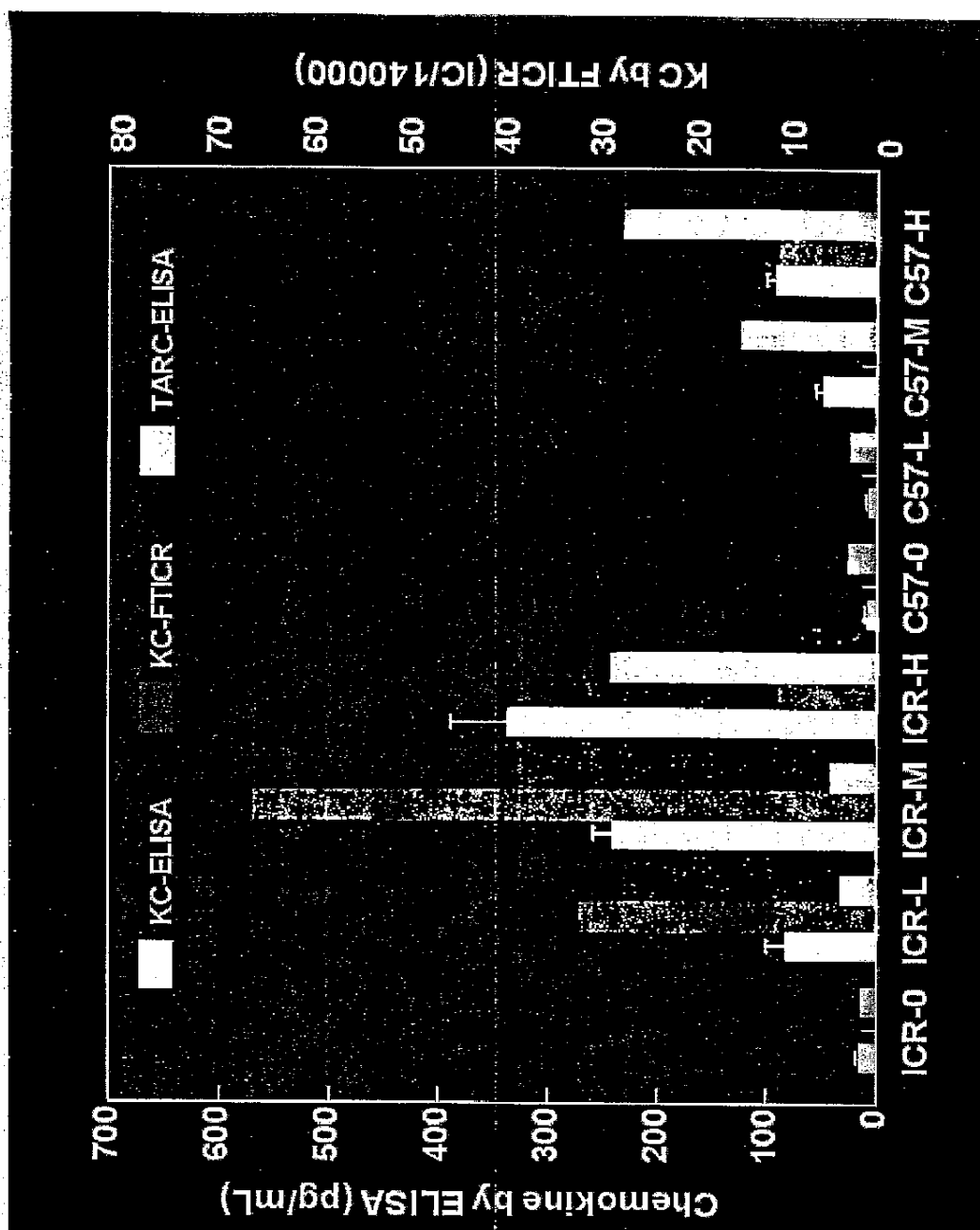
KC peptide (AMT)  
detected as shown.

Four samples were  
below detection  
limits.

Not precisely the  
same sample

So what did we  
learn?

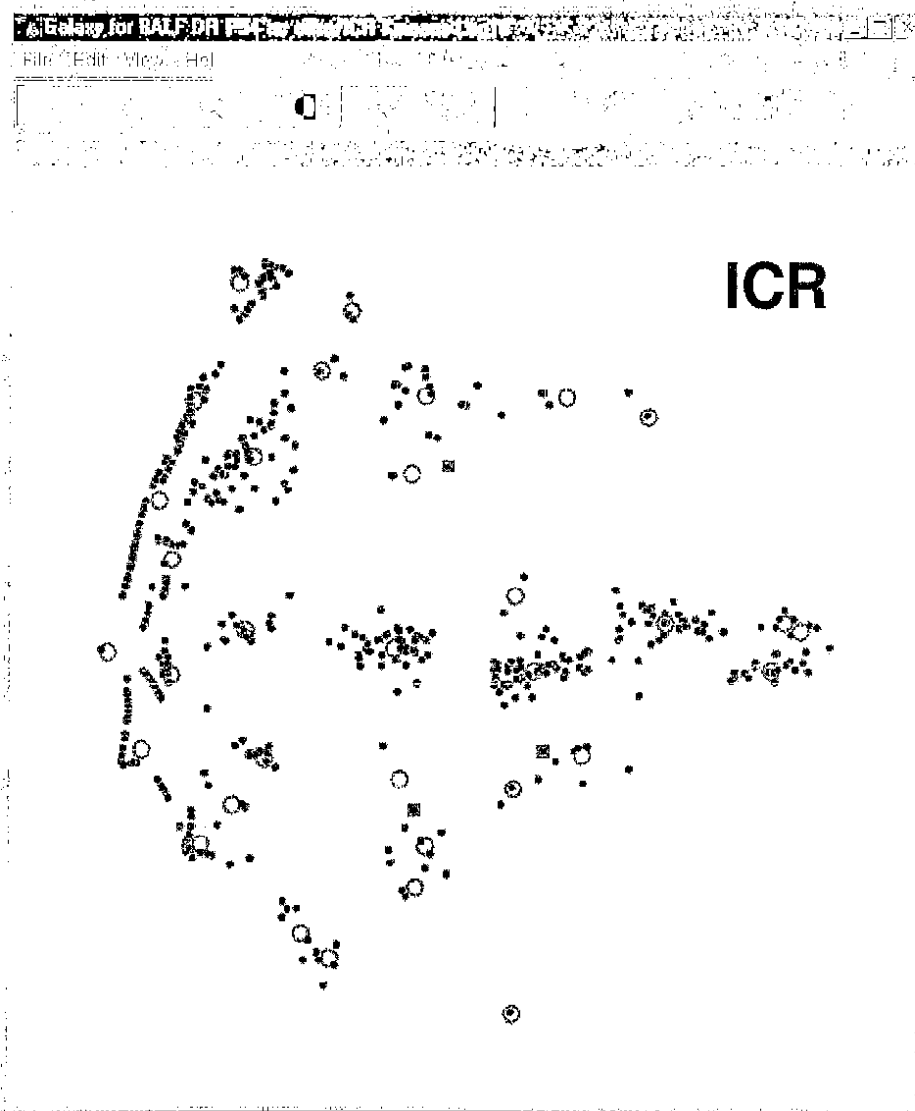
Battelle





# Galaxy View: Looking at the data

- Two-D projection of higher dimensional data (x and y axes have no units or dimensions)
- Used to evaluate distribution of data to compare data sets and subsets at a high level
- Each of 1154 proteins identified is represented by a single blue square
- Closely related data cluster together, while dissimilar data are separated by larger distances
- The small circles define the centroid for each of the 34 clusters or groups of proteins that have similar behavior across the columns of data
- Many clustering options and algorithms available

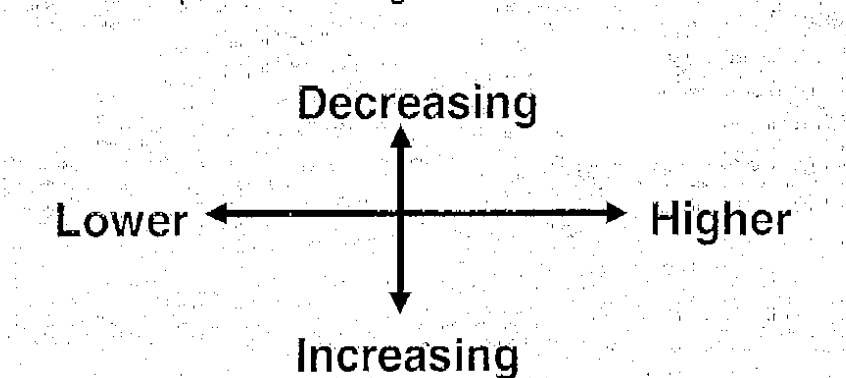


Battelle

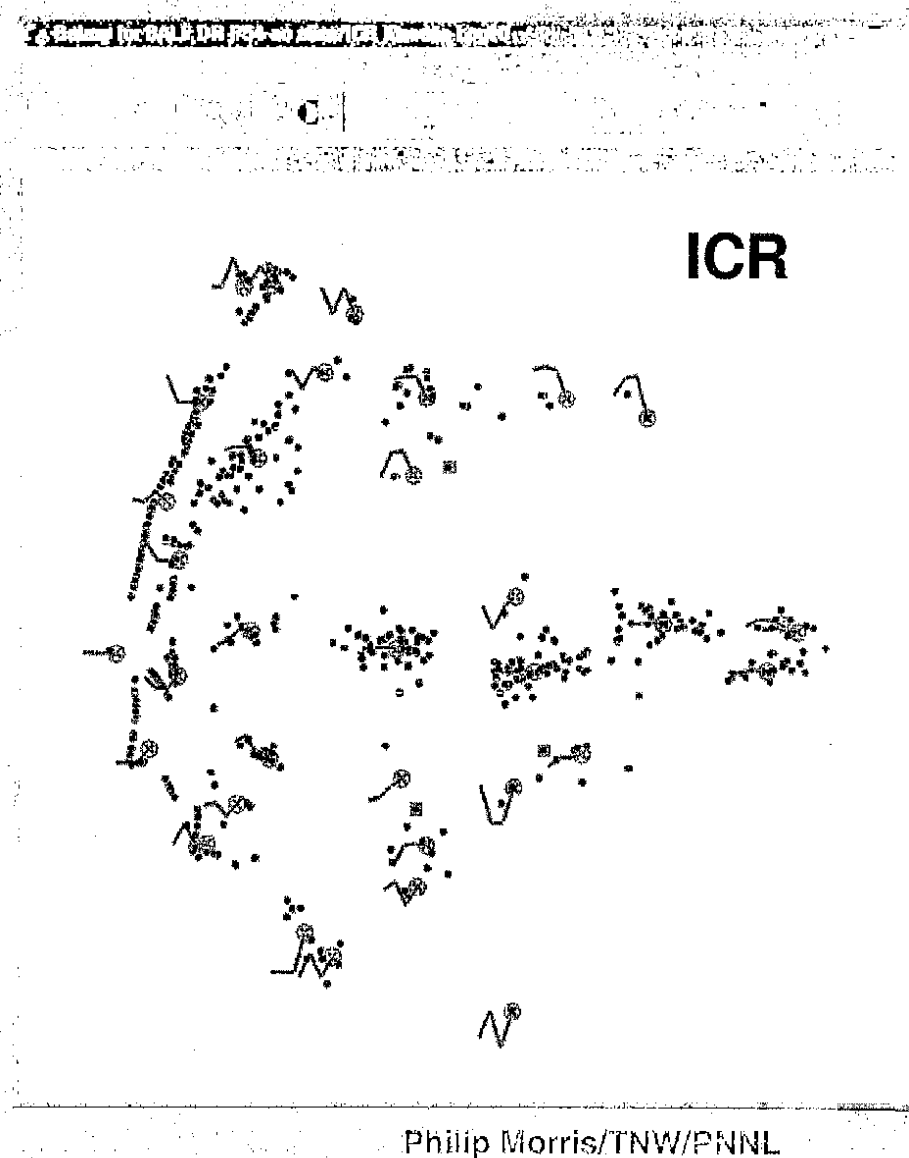
Philip Morris/TNW/PNNL

38

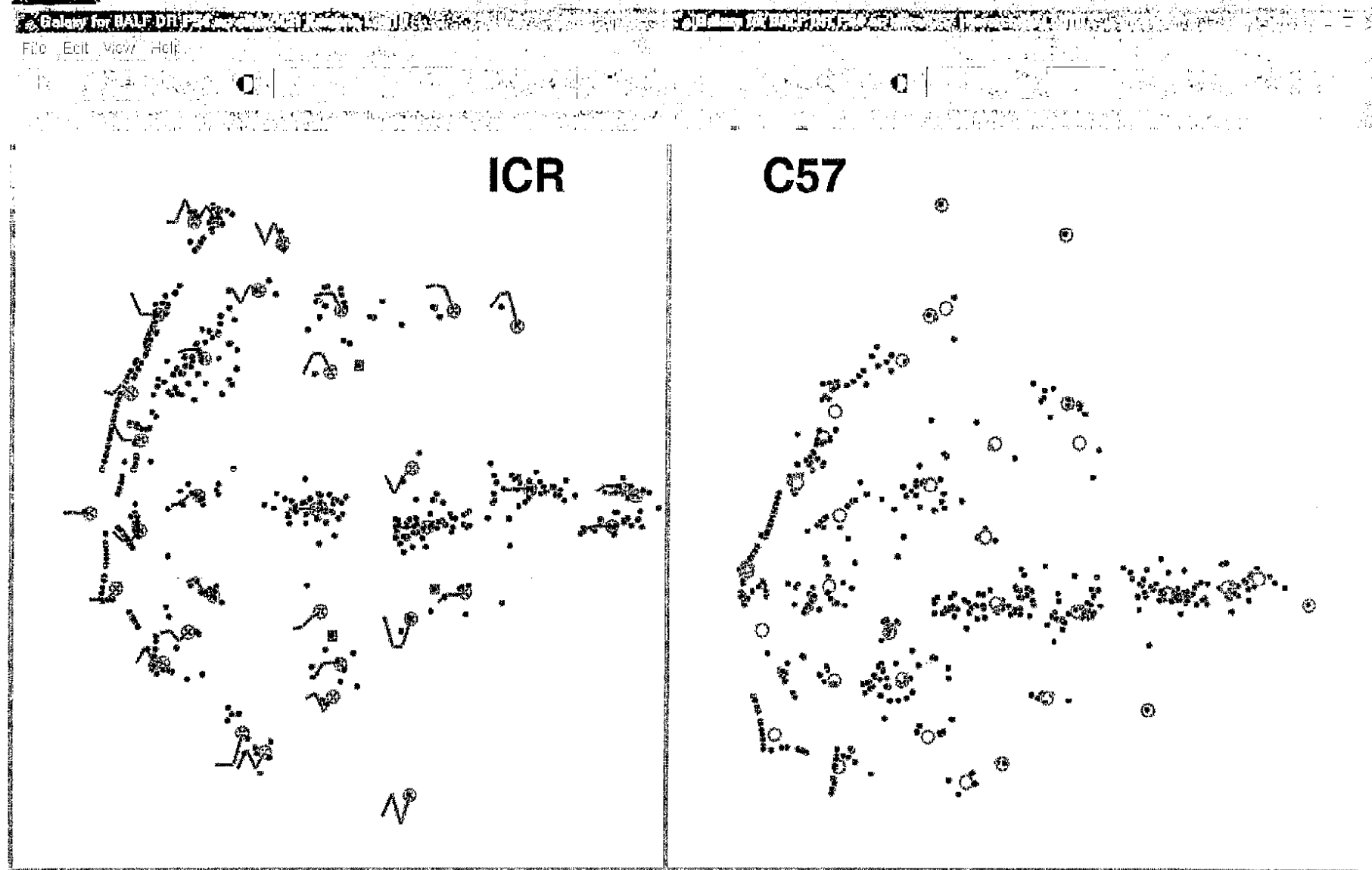
- Used to evaluate the entire galaxy of data and to identify clusters of further interest
- Shows the shape of the data in the individual clusters and is equivalent to the plot of the mean response shown in 'DataPlot'
- When the data may be arranged in the order of increasing dose or time, the interpretation of miniplots is straightforward.



Batelle



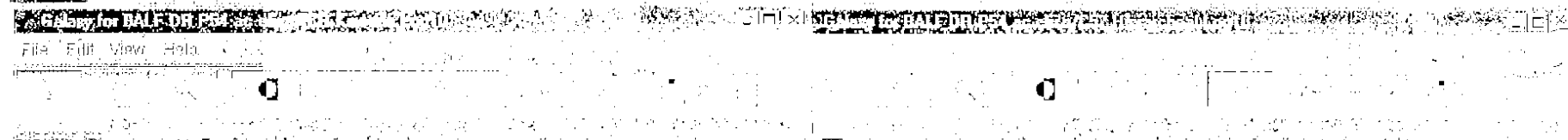
# Galaxy View ICR vs. C57



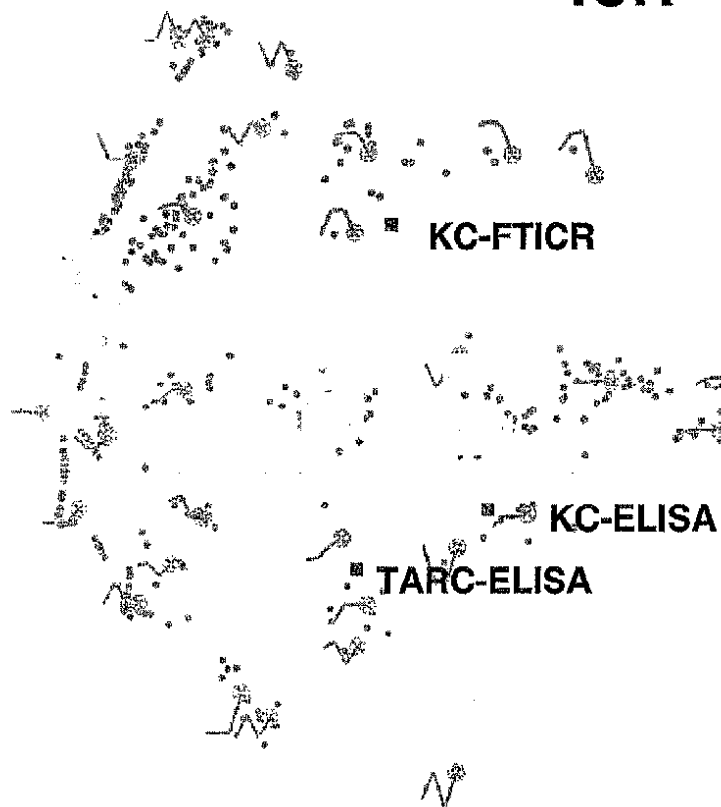
PM3001148287

# What about KC?

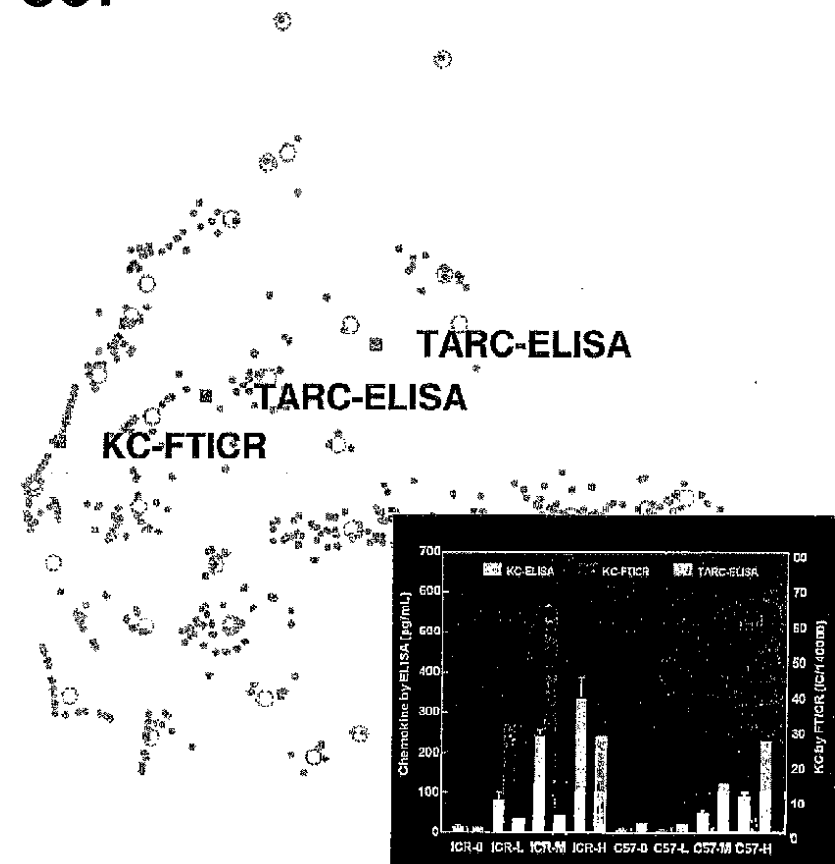
(KC & TARC ELISA scaled)



**ICR**



**C57**



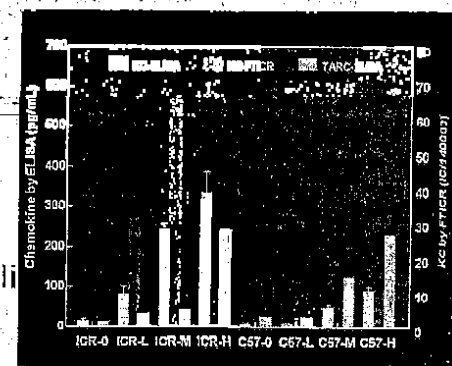
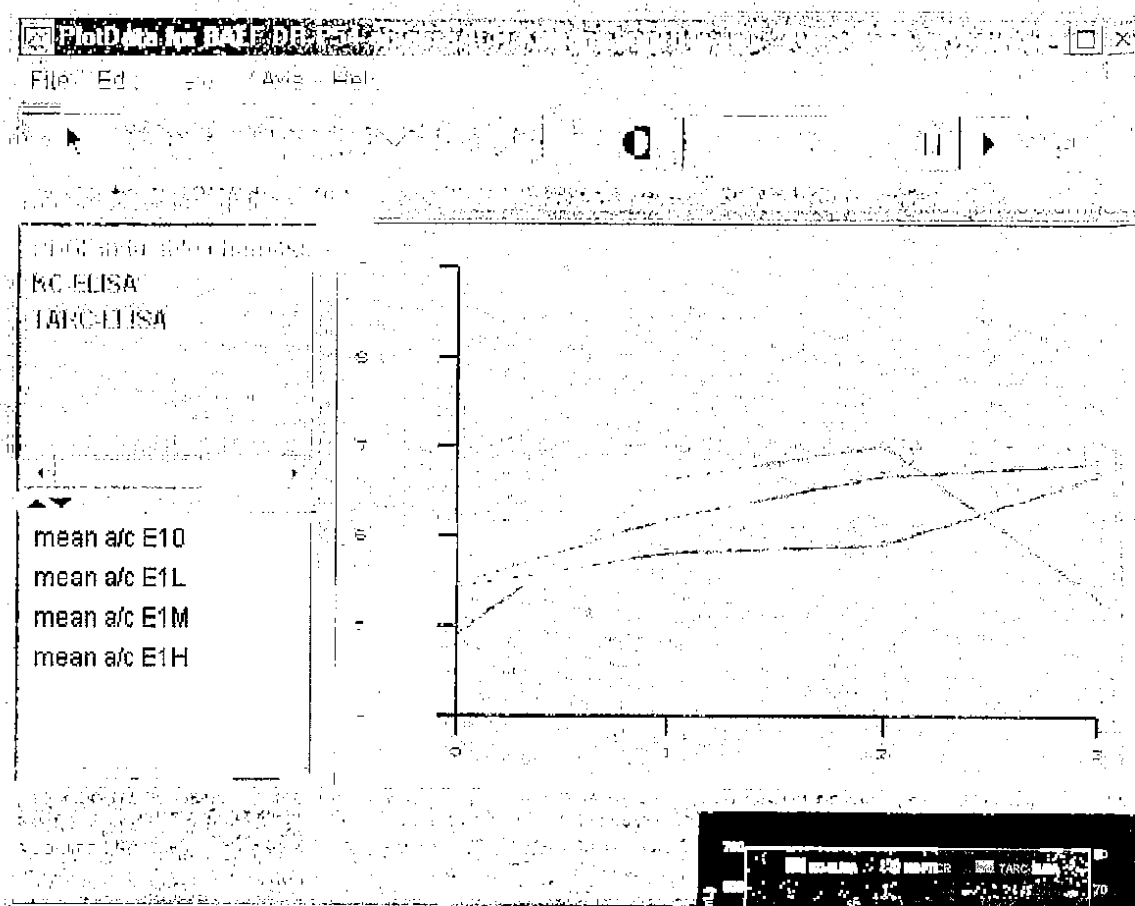
PM3001148288

# DataPlot View: What other proteins have dose-response patterns like KC?

May display a single cluster, multiple clusters, and individual or selected proteins.

The X-axis is the four data columns identified by their code shown in the lower left panel.

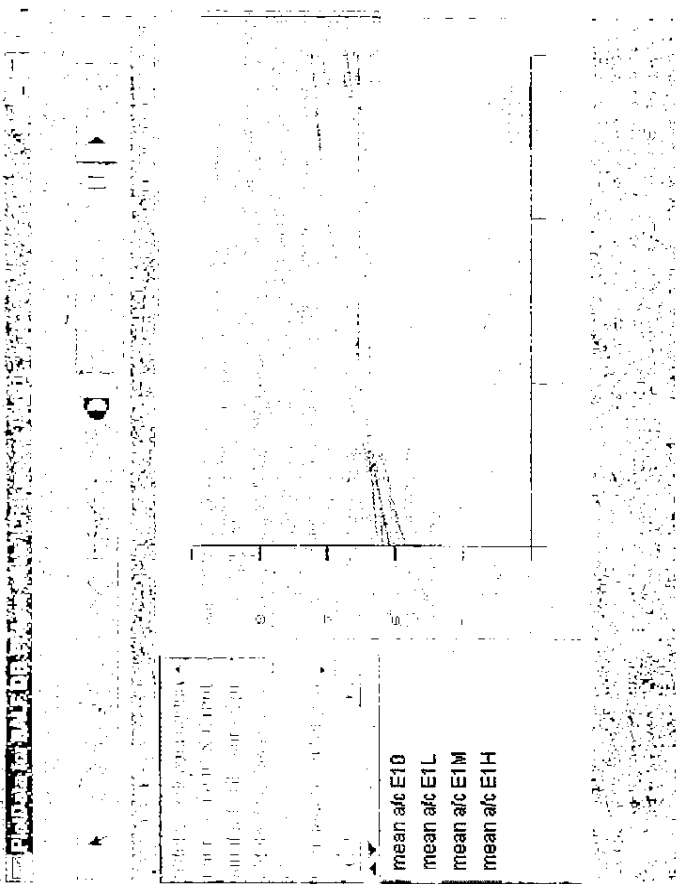
The upper left panel contains the names of the selected proteins that are color matched to the lines on the plot.



Battelle

Phili

# What proteins respond similar to KC-ELISA?



Battelle

PM3001148290

# Proteins similar to KC-ELISA in ICR?

- KC-ELISA
- aldehyde dehydrogenase family 1; subfamily A1;
- aldehyde dehydrogenase family 1; subfamily A7;
- carbonic anhydrase 2; CA II
- ceruloplasmin
- chemokine (C-X-C motif) ligand 15; small inducible cytokine subfamily B; member 15
- kininogen; H-kininogen; L-kininogen
- similar to ribosomal protein L13; cytosolic [validated] - rat
- transient receptor potential cation channel; subfamily V; member 6;

Battelle

Philip Morris/TNW/CPM

69

# Summary:

## Sample Preparation:

- Mean ion-intensity provides good measure of relative abundance of individual peptides/proteins across experimental groups
- Lack of dominant proteins simplifies preparation of BALF for MS analysis

## Number of Proteins identified in BALF increased ten fold:

- Several hundred proteins are reproducibly identified in BALF significantly increasing knowledge of this fluid.
- BALF contain many proteins from different functional classes including inflammatory, secreted, and cellular proteins.

## Strain- and exposure-dependent differences in protein composition.

- Numerous strain- and exposure-dependent differences observed
  - ~80% of proteins do not change in abundance with treatment
  - ~8% decrease; 4% increase; 8% have complex responses to increasing dose
- Pooled samples precludes statistical analysis

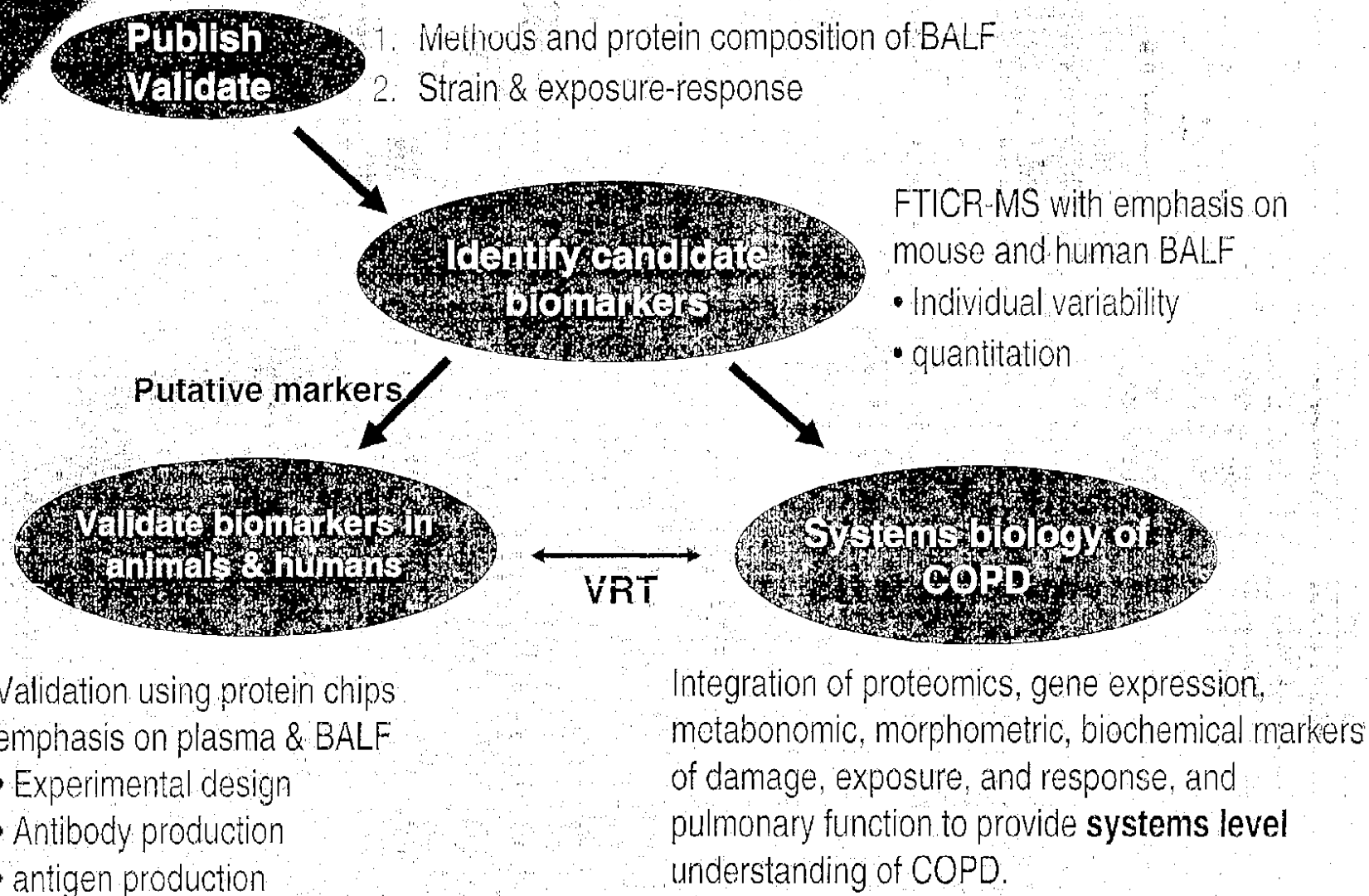
Battelle

Philip Morris/TNW/PN0011

PM3001148292



# Future COPD



Battelle

Philip Morris/TNW/PNNL

PM3001148293

# Acknowledgements

## Sources of Funding

Philip Morris (Willie McKinney)  
Battelle (Robert Miller, Mark Gritz, Bruce  
Westerberg)

## Biology

Josh Adkins, David Springer, Susan Varnum,  
Richard Zangar

## Mass Spectroscopy

Richard Smith, Ron Moore, David Wunschel

## Bioinformatics/Data processing

Gordon Anderson, Ken Auberry, Nikola Tolic

## In-Life/Toxicology

K. Monica Lee

Battelle



Philip Morris/TNW/PNNL